

UNIVERSIDADE FEDERAL DO PARANÁ

IVAN MOYSES GUIMARÃES

**PREVISÃO DO PRODUTO INTERNO BRUTO MUNICIPAL VIA EMPREGO
CELETISTA: UMA PROPOSTA METODOLÓGICA**

CURITIBA

2018

IVAN MOYSES GUIMARÃES

**PREVISÃO DO PRODUTO INTERNO BRUTO MUNICIPAL VIA EMPREGO
CELETISTA: UMA PROPOSTA METODOLÓGICA**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Métodos Numéricos em Engenharia pelo Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática, dos Setores de Tecnologia e de Ciências Exatas, da Universidade Federal do Paraná.

Orientador: Prof. Dr. Anselmo Chaves Neto
Coorientadora: Profa. Dra. Sachiko Araki Lira

CURITIBA

2018

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS/UFPR
BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

G963p

Guimarães, Ivan Moyses

Previsão do produto interno bruto municipal via emprego celetista: uma proposta metodológica / Ivan Moyses Guimarães. – Curitiba, 2018.
101 p. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Tecnologia e de Ciências Exatas, Programa de Pós-Graduação em Métodos Numéricos em Engenharia, 2018.

Orientador: Anselmo Chaves Neto.

Coorientadora: Sachiko Araki Lira.

1. Produto interno bruto . 2. Emprego celetista. 3. Regressão. 4. Séries temporais. I. Universidade Federal do Paraná. II. Chaves Neto, Anselmo. III. Lira, Sachiko Araki. IV. Título.

CDD: 339.32



MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MÉTODOS NUMÉRICOS
EM ENGENHARIA

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MÉTODOS NUMÉRICOS EM ENGENHARIA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **IVAN MOYSES GUIMARÃES** intitulada: **PREVISÃO DO PRODUTO INTERNO BRUTO MUNICIPAL VIA EMPREGO CELETISTA: UMA PROPOSTA METODOLÓGICA**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 09 de Março de 2018.

ANSELMO CHAVES NETO

Presidente da Banca Examinadora (UFPR)

JAIR MENDES MARQUES

Avaliador Interno (UFPR)

ARMANDO VAZ SAMPAIO

Avaliador Externo (UFPR)

AGRADECIMENTOS

Agradeço a Deus por ter me concedido a benção de ingressar no mestrado, me manter e ainda o concluir. Por ter me concedido saúde, inteligência e capacidade a ponto de vencer mais esta etapa. Por ter colocado pessoas fundamentais na minha vida que me auxiliaram nesta pesquisa. Agradeço por Ele ter cuidado de mim até agora, sabendo que Ele ainda cuidará.

Agradeço a minha mentora Sachiko Araki Lira pela orientação, disposição, pela cobrança e por toda a ajuda, pois sem ela esta pesquisa não seria possível.

Agradeço ao meu orientador Anselmo Chaves Neto por se mostrar aberto a desempenhar tal função, depositando total confiança na pesquisa desenvolvida.

Agradeço ao Guilherme Amorin, do IPARDES, pela disposição em esclarecer dúvidas econômicas sempre que solicitado.

Agradeço ao Nivaldo Aparecido Minervi por ter cedido o início da função *best.arima()*, além de conselhos sobre séries temporais.

Agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de bolsa de mestrado para a realização deste estudo.

Agradeço aos meus pais, Timoteo e Olga, e irmãos, Isaque e Paulo, pelo apoio a mais uma etapa na minha vida, assim como em todas as outras. Por terem me animado e fortalecido sempre que necessário, e até mesmo quando não era necessário.

Agradeço aos meus colegas do PPGMNE, Maicon Malacarne, Mauren Micalichen, Nicole Rozin, Lucas Lamy, Bruno e Inajara Rutyna, Michely de Oliveira, Igor Corazza, Diego Metz, que estiveram comigo em todos os momentos, sendo eles tristes ou alegres, se mostrando grandes e verdadeiros amigos, podendo contar com cada um deles não só durante a pesquisa, mas, também, durante a vida.

Agradeço aos meus amigos Eduardo e Daiana Graebim, por serem tão compreensivos no estágio final da escrita da dissertação.

Enfim, agradeço a todos que direta ou indiretamente contribuíram para a conclusão deste trabalho.

Saber para prever, prever para controlar.

Auguste Comte

RESUMO

Um dos índices de crescimento econômico mais conhecido e utilizado para compreender a situação em que se encontra a economia de uma certa região é o Produto Interno Bruto (PIB). Sendo assim, conhecer o comportamento do PIB é de fundamental importância para entender o crescimento, ou decréscimo, da região a ser analisada. A partir da compreensão do desempenho do PIB pode-se fazer uma projeção desse indicador. Entretanto, devido a curta série temporal mensal existente do PIB municipal, a projeção direta deste índice se torna difícil. Sendo assim, este trabalho apresenta uma alternativa para a previsão do PIB municipal paranaense a partir de outra variável fortemente correlacionada com o PIB, no caso, o emprego celetista, analisada através da correlação de Spearman. Após definida a variável emprego celetista, realizou-se uma Análise Exploratória de Dados Espaciais (AEDE) para identificar *clusters* espaciais dos municípios do Paraná utilizando o Índice Local de Associação Espacial. Os municípios estudados na pesquisa, foram identificados através de *clusters* de municípios que apresentaram PIB elevado, e as médias das cidades vizinhas também elevadas, classificadas como H-H (alto-alto) para o ano de 2013, tendo sido o ano que apresentou maior PIB a preços constantes do estado do Paraná, no período de 2002 a 2014. A fim de encontrar uma relação funcional entre as duas variáveis estudadas, esta pesquisa fez uso de modelos de regressão, obtendo valores de R^2 ajustados bem próximos de 1, considerando o período já mencionado, devido os dados do PIB municipal estarem disponíveis apenas neste intervalo de tempo. Através de uma Análise de Séries Temporais, ajustou-se modelos de previsão para o emprego celetista a fim de usar estas previsões na relação funcional e assim prever o valor do PIB municipal.

Palavras-chave: PIB. Emprego Celetista. Regressão. Séries Temporais.

ABSTRACT

One of the best-known economic growth indices used to understand the situation of a region's economy is the Gross Domestic Product (GDP). Therefore, knowing the behavior of the GDP is of fundamental importance to understand the growth, or decrease, of the region to be analyzed. From the understanding of the performance of the GDP one can make a projection of this indicator. However, due to the short monthly time series of municipal GDP, the direct projection of this index becomes difficult. Thus, this paper presents an alternative for the prediction of the municipal GDP of Paraná from another variable strongly correlated with the GDP, in this case, the bargaining employment, analyzed through the Spearman correlation. After defining the variable occupation, an Exploratory Spatial Data Analysis (ESDA) was carried out to identify spatial clusters of the municipalities of Paraná using the Local Spatial Association Index. The municipalities studied in the research were identified through clusters of municipalities that presented high GDP, and the averages of neighboring cities also high, classified as HH (high-high) for the year 2013, being the year that presented the highest GDP a constant price of the state of Paraná, between 2002 and 2014. In order to find a functional relationship between the two variables studied, this research made use of regression models, obtaining adjusted R^2 values very close to 1, considering the period already municipal GDP data is available only in this time interval. Through a Time Series Analysis, forecast models were fitted for the job market in order to use these predictions in the functional relation and thus to predict the value of municipal GDP.

Key-words: GDP. Formal employment. Regression. Time Series

LISTA DE FIGURAS

FIGURA 1 – MAPA DE CLUSTER LISA DO PIB A PREÇOS CONSTANTES (2014 = 100) DOS MUNICÍPIOS PARANAENSES – 2013.....	44
FIGURA 2 – FLUXOGRAMA DA METODOLOGIA PROPOSTA.....	50

LISTA DE GRÁFICOS

GRÁFICO 1	–	VARIAÇÃO EM VOLUME DO PIB TRIMESTRAL, COM AJUSTE SAZONAL - BRASIL – 1º TRIM. 2012 – 1º TRIM. 2016.....	21
GRÁFICO 2	–	SÉRIE TEMPORAL DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CURITIBA - JAN. 2007- DEZ. 2014.....	55
GRÁFICO 3	–	AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO SARIMA(0,2,1)X(2,2,0) ₁₂ , AJUSTADO PARA O MUNICÍPIO DE CURITIBA.....	57
GRÁFICO 4	–	P-VALORES DAS 25 PRIMEIRAS AUTO CORRELAÇÕES DOS RESÍDUOS DO MODELO SARIMA(0,2,1)X(2,2,0) ₁₂ , AJUSTADO PARA O MUNICÍPIO DE CURITIBA.....	57
GRÁFICO 5	–	PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO SARIMA(0,2,1)X(2,2,0) ₁₂ , AJUSTADO PARA O MUNICÍPIO DE CURITIBA.....	58
GRÁFICO 6	–	PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CURITIBA PELO MODELO ARIMA(0,2,1)X(2,2,0) ₁₂	59
GRÁFICO 7	–	SÉRIE TEMPORAL DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ARAUCÁRIA - JAN. 2007- DEZ. 2014.....	60
GRÁFICO 8	–	PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ARAUCÁRIA PELO MODELO MMS(1)....	61
GRÁFICO A.3.1	–	AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ.....	74
GRÁFICO A.3.2	–	PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ.....	75
GRÁFICO A.3.3	–	PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ PELO MODELO DE PREVISÃO AJUSTADO.....	75

GRÁFICO A.4.1	– AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ARAPONGAS.....	76
GRÁFICO A.4.2	– PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ARAPONGAS.....	77
GRÁFICO A.4.3	– PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ARAPONGAS PELO MODELO DE PREVISÃO AJUSTADO.	77
GRÁFICO A.5.1	– AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMBÉ	78
GRÁFICO A.5.2	– PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMBÉ.....	79
GRÁFICO A.5.3	– PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CAMBÉ PELO MODELO DE PREVISÃO AJUSTADO.	79
GRÁFICO A.6.1	– AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMPO LARGO.....	80
GRÁFICO A.6.2	– PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMPO LARGO.....	81
GRÁFICO A.6.3	– PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CAMPO LARGO PELO MODELO DE PREVISÃO AJUSTADO.	81
GRÁFICO A.7.1	– AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CARAMBEÍ.....	82
GRÁFICO A.7.2	– PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CARAMBEÍ.....	83

GRÁFICO A.7.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CARAMBEÍ PELO MODELO DE PREVISÃO AJUSTADO	83
GRÁFICO A.8.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE COLOMBO	84
GRÁFICO A.8.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE COLOMBO	85
GRÁFICO A.8.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE COLOMBO PELO MODELO DE PREVISÃO AJUSTADO.	85
GRÁFICO A.9.1 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE FAZENDA RIO GRANDE PELO MODELO MMS(1).....	86
GRÁFICO A.10.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE IBIPORÃ	87
GRÁFICO A.10.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE IBIPORÃ	88
GRÁFICO A.10.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE IBIPORÃ PELO MODELO DE PREVISÃO AJUSTADO.	88
GRÁFICO A.11.1 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE FAZENDA RIO GRANDE PELO MODELO MMS(1).....	89
GRÁFICO A.12.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE PIRAQUARA	90
GRÁFICO A.12.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE PIRAQUARA.....	91

GRÁFICO A.12.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE PIRAQUARA PELO MODELO DE PREVISÃO AJUSTADO	91
GRÁFICO A.13.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ROLÂNDIA	92
GRÁFICO A.13.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ROLÂNDIA	93
GRÁFICO A.13.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ROLÂNDIA PELO MODELO DE PREVISÃO AJUSTADO.	93
GRÁFICO A.14.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS	94
GRÁFICO A.14.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS.....	95
GRÁFICO A.14.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS PELO MODELO DE PREVISÃO AJUSTADO.....	95
GRÁFICO A.15.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SARANDI.....	96
GRÁFICO A.15.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SARANDI.....	97
GRÁFICO A.15.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE SARANDI PELO MODELO DE PREVISÃO AJUSTADO.	97

LISTA DE TABELAS

TABELA 1	–	MUNICÍPIOS LIMÍTROFES, INDICADOR PADRONIZADO, MÉDIA DO INDICADOR PADRONIZADO DOS MUNICÍPIOS LIMÍTROFES E CLASSIFICAÇÃO, SEGUNDO ALGUNS MUNICÍPIOS, PARANÁ – 2013.....	45
TABELA 2	–	DADOS DO PIB, EMPREGO CELETISTA E DUMMY PARA O MUNICÍPIO DE ARAUCÁRIA – 2002-2014	48
TABELA 3	–	COEFICIENTES ESTIMADOS, ESTATÍSTICA T, P-VALOR E R^2 AJUSTADO, DE MODELOS DE REGRESSÃO ESTIMADOS PARA 2 MUNICÍPIOS SELECIONADOS DO ESTADO DO PARANÁ	51
TABELA 4	–	COEFICIENTES ESTIMADOS, ESTATÍSTICA T, P-VALOR E R^2 AJUSTADO, DE MODELOS DE REGRESSÃO ESTIMADOS PARA 13 MUNICÍPIOS SELECIONADOS DO ESTADO DO PARANÁ	52
TABELA 5	–	RESULTADOS DOS TESTES DAS ANÁLISES DOS RESÍDUOS DE MODELOS DE REGRESSÃO ESTIMADOS PARA 15 MUNICÍPIOS SELECIONADOS DO ESTADO DO PARANÁ.....	53
TABELA 6	–	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO_SARIMA(0,2,1)X(2,2,0) ₁₂ , AJUSTADO PARA O MUNICÍPIO DE CURITIBA	56
TABELA 7	–	MODELOS DE PREVISÃO AJUSTADO PARA 15 MUNICÍPIOS PARANAENSES SELECIONADOS E O CRITÉRIO AIC.	60
TABELA 8	–	PREVISÃO DO PIB MUNICIPAL UTILIZANDO A METODOLOGIA PROPOSTA EM COMPARAÇÃO COM O VALOR DO PIB A PREÇOS CONSTANTES (2014=100) (R\$ 1.000).....	63
TABELA A.1.1	–	CORRELAÇÃO DE SPEARMAN ENTRE PIB A PREÇOS CONSTANTES (2014 = 100) DOS MUNICÍPIOS DO PARANÁ SEGUNDO ALGUMAS VARIÁVEIS	72

TABELA A.2.1 –	RESULTADOS DO TESTE DE JARQUE-BERA E LJUNGBOX PARA OS 15 RESÍDUOS DOS MODELOS DE PREVISÃO AJUSTADOS NO ESTÁGIO III.....	73
TABELA A.3.1 –	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ	74
TABELA A.4.1 –	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ARAPONGAS	76
TABELA A.5.1 –	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMBÉ	78
TABELA A.6.1 –	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMPO LARGO	80
TABELA A.7.1 –	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CARAMBEÍ	82
TABELA A.8.1 –	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE COLOMBO.....	84
TABELA A.10.1–	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE IBIPORÃ	87
TABELA A.12.1–	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE PIRAQUARA.....	90
TABELA A.13.1–	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ROLÂNDIA.....	92
TABELA A.14.1–	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS	94

TABELA A.15.1–	COEFICIENTES ESTIMADOS, ERRO PADRÃO E P- VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SARANDI	96
----------------	---	----

SUMÁRIO

1	INTRODUÇÃO.....	18
1.1	OBJETIVOS	20
1.1.1	Objetivo Geral	20
1.1.2	Objetivos Específicos.....	20
1.2	JUSTIFICATIVA.....	20
1.3	ESTRUTURA DO TRABALHO	22
2	REVISÃO DE LITERATURA	23
2.1	PRODUTO INTERNO BRUTO	23
2.1.1	PIB nominal.....	25
2.1.2	PIB real	25
2.2	DEFLATOR DO PIB.....	25
2.3	ÍNDICES DE EMPREGO	26
2.4	CORRELAÇÃO DE SPEARMAN	28
2.5	ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS (AEDE)	29
2.5.1	Índice Local De Associação Espacial (LISA)	29
2.6	ANÁLISE DE REGRESSÃO	30
2.7	ANÁLISE DE SÉRIES TEMPORAIS.....	32
3	MATERIAL E MÉTODOS	42
3.1	MATERIAL	42
3.2	MÉTODOS.....	43
3.2.1	Estágio I – Tratamento e seleção de dados.....	43
3.2.2	Estágio II – Análise de Regressão Múltipla.....	47
3.2.3	Estágio III – Análise de Séries Temporais	49
4	APRESENTAÇÃO DOS RESULTADOS.....	51
4.1	RESULTADOS DA ANÁLISE DE REGRESSÃO	51
4.2	RESULTADOS DA ANÁLISE DE SÉRIES TEMPORAIS.....	55
4.3	PREVISÃO DO PIB DE CURITIBA PARA O ANO DE 2015.....	62
4.4	PREVISÃO DO PIB PARA OS DEMAIS MUNICÍPIOS SELECIONADOS.....	62
5	CONCLUSÃO	64
6	CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES.....	66
	REFERÊNCIAS	67

APÊNDICE 1 – CORRELAÇÃO DE SPEARMAN ENTRE PIB DOS MUNICÍPIOS E ALGUMAS OUTRAS VARIÁVEIS	72
APÊNDICE 2 – RESULTADOS DOS TESTES DE NORMALIDADE E LJUNG-BOX DOS RESÍDUOS PARA OS MODELOS DE PREVISÃO AJUSTADOS	73
APÊNDICE 3 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ.....	74
APÊNDICE 4 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE ARAPONGAS.....	76
APÊNDICE 5 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE CAMBÉ.....	78
APÊNDICE 6 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE CAMPO LARGO.....	80
APÊNDICE 7 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE CARAMBEÍ.....	82
APÊNDICE 8 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE COLOMBO	84
APÊNDICE 9 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE FAZENDA RIO GRANDE	86
APÊNDICE 10 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE IBIPORÃ.....	87
APÊNDICE 11 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE PINHAIS	89
APÊNDICE 12 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE PIRAQUARA	90
APÊNDICE 13 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE ROLÂNDIA.....	92
APÊNDICE 14 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS	94
APÊNDICE 15 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE SARANDI.....	96
APÊNDICE 16 – FUNÇÃO BEST.ARIMA() UTILIZANDO O SOFTWARE R CORE TEAM	98

1 INTRODUÇÃO

As discussões sobre o crescimento da economia se deram, inicialmente, de forma mais filosófica com estudiosos como Adam Smith, David Ricardo, Thomas Malthus, John Stuart Mill, entre tantos outros que analisavam e inferiam sobre as tendências do Mercado com base apenas nas observações de suas épocas. Entretanto, com o passar dos anos, os estudos do crescimento econômico têm obtido maiores respaldos devido às séries de dados que foram armazenados no decorrer dos anos e, ainda, com avanços nas teorias estatísticas, teorias de Análise de Dados, surgimento da econometria e outras ferramentas para estudar o crescimento econômico (GARCÍA; URDANIVIA, 2000).

Nos últimos anos, muitos meios de comunicação, além de pessoas físicas e políticas do cotidiano dos brasileiros, têm falado sobre a crise econômica que o Brasil está passando, levando muitos cidadãos brasileiros e até mesmo investidores estrangeiros a uma preocupação com a economia do país. Devido a isso, o estudo do Produto Interno Bruto (PIB) torna-se de suma importância.

Segundo Silva et al. (2015), o PIB é o principal indicador da atividade econômica, tendo por objetivo detalhar o valor da produção em um período determinado. O valor do PIB de um país ou região representa a produção de todos os setores produtores da economia (empresas públicas e privadas produtoras de bens e prestadoras de serviços, trabalhadores autônomos, governo, etc.), em um período estipulado, avaliados a preço de mercado (FEIJÓ; RAMOS, 2013).

Guimarães e Lira (2017b) destacam que o crescimento de uma região a nível de estado, pode ser maior ou menor que o PIB do país. Os autores citam que “uma análise sobre a tendência do PIB regional torna-se necessária, mesmo a curto prazo” (GUIMARÃES; LIRA, 2017b, p. 2).

A importância de se individualizar o estudo por região, desagregando a nível de municípios, deve-se às diferentes características que cada um apresenta, seja no contexto econômico, geográfico, político ou social. Essas particularidades influenciam de formas diferenciadas nos municípios.

Segundo Hobsbawm (2013, p.40) “toda a previsão sobre o mundo real repousa, em grande parte, em algum tipo de inferência sobre o futuro a partir daquilo que aconteceu no passado”. Sendo assim, podemos definir que uma previsão pode ser classificada como uma estimativa quantitativa sobre eventos futuros a partir de

informações de períodos passados. Desta forma, técnicas de previsão qualitativas e quantitativas são desenvolvidas (STEFFEN, 2010).

Quando se há poucos dados, usam-se técnicas qualitativas de previsão, sendo necessário consultar especialistas no assunto para analisarem os dados que se têm, e, então, seus pareceres costumam ser a palavra final da previsão.

Quando a série de dados apresenta um histórico maior, usam-se técnicas quantitativas de previsão, as quais baseiam-se em aplicar diferentes métodos estatísticos com o objetivo de extrapolar os dados e obter prováveis valores futuros. Nesse sentido, esta pesquisa apoia-se na pesquisa quantitativa, através de métodos estatísticos, a fim de prever os valores do PIB a partir de sequências de valores passados.

Sendo assim, este trabalho analisa os dados do PIB municipal anual de 15 municípios do estado do Paraná, utilizando-se dos valores constantes do PIB para o período de 2002 a 2014, a fim de poder fazer uma previsão do PIB em 2015 para estes municípios. Entretanto, como o número de dados é muito pequeno, busca-se uma alternativa para fazer esta previsão.

A alternativa encontrada é de utilizar uma variável que apresenta uma correlação de Spearman muito forte com o PIB municipal, obtendo uma série de dados maior que a do PIB, a fim de ser feita uma inferência para os valores do PIB municipal a partir de uma relação funcional entre o PIB e a previsão dos valores desta variável para o ano de 2015.

A relação funcional entre PIB e a variável fortemente correlacionada com ele, é desenvolvida através da Análise de Regressão Linear Múltipla, sendo escolhidos os modelos com valores altos do R^2 ajustado.

Para prever os valores da variável altamente correlacionada com o PIB, este trabalho utiliza modelos de previsão Autorregressivos, Integrados e de Médias Móveis (ARIMA) com sazonalidade, popularmente difundido por Box e Jenkins (BOX; JENKINS; REINSEL, 1994).

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Propor uma nova metodologia para obtenção da previsão do PIB municipal anual, a partir de outra variável altamente correlacionada com o mesmo, mas que tenha uma série temporal de dados com período de tempo maior, devido ao fato da série temporal do PIB municipal ser curta.

1.1.2 Objetivos Específicos

Os objetivos específicos do trabalho são:

- a) definir qual das variáveis é mais correlacionada com o PIB municipal a preços constantes: rendimento médio mensal dos trabalhadores paranaenses, consumo de energia elétrica total, industrial e comercial e emprego celetista;
- b) identificar as aglomerações (*clusters*) espaciais do PIB real (preços constantes) dos municípios do estado do Paraná, utilizando as técnicas de análise de dados espaciais de área, para o ano selecionado de 2013;
- c) ajustar modelos de regressão para encontrar uma relação funcional entre PIB e a variável altamente correlacionada obtida no objetivo “a”;
- d) ajustar modelos de previsão para a variável mais correlacionada com o PIB dos municípios que apresentaram correlação positiva nas aglomerações espaciais obtidas no objetivo “b”.

1.2 JUSTIFICATIVA

Com vista ao país estar passando por dificuldades financeiras, este trabalho propõe uma metodologia de projeção do PIB municipal, para apontar possibilidades de investimento respaldadas em técnicas estatísticas, eximindo-se da subjetividade de avaliação de um especialista na área econômica, trabalhando apenas em função dos dados obtidos.

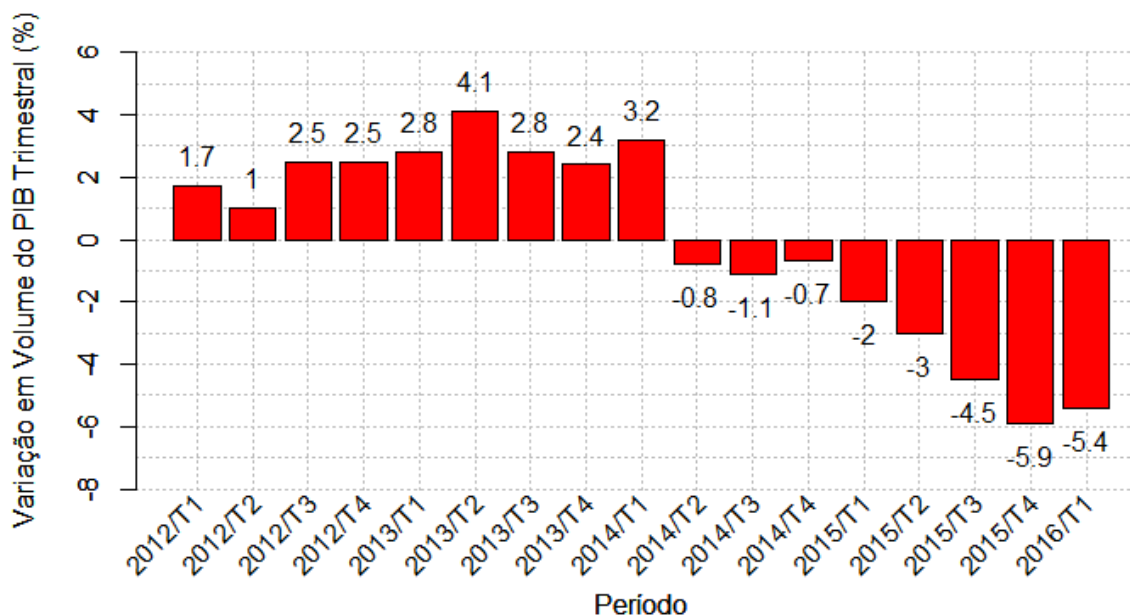
Informações confiáveis são necessárias para atrair investimentos para os municípios paranaenses. Também, para servir de apoio para elaboração de políticas públicas.

Segundo Bacha (2017, p. 23),

desde 2014, o Brasil atravessa uma das mais longas e profundas recessões de sua história. Entre 2013 a 2016, o Produto Interno Bruto (PIB) per capita acumulou uma queda de 9%. Dito de outro modo, os brasileiros chegam a 2017 em média 9% mais pobres do que estavam no início de 2014.

Através do site do instituto brasileiro de geografia e estatística (IBGE) pode-se obter a série estatística da variação em volume do PIB trimestral brasileiro desde 1996 até o primeiro trimestre de 2016. O GRÁFICO 1 abaixo, mostra a série da variação em volume do PIB a partir do primeiro trimestre de 2012 até o primeiro trimestre de 2016.

GRÁFICO 1 – VARIAÇÃO EM VOLUME DO PIB TRIMESTRAL, COM AJUSTE SAZONAL - BRASIL
– 1º TRIM. 2012 – 1º TRIM. 2016



FONTE: IBGE. Diretoria de Pesquisas Coordenação de Contas Nacionais (2017)

Nota-se que a partir do segundo trimestre de 2014, a variação em volume do PIB no Brasil, vem apresentando taxas negativas. Assim, a datar do terceiro trimestre de 2014, principiou-se o período de recessão econômica no Brasil, devido ao fato de ser o segundo trimestre consecutivo de encolhimento do PIB.

Entretanto, observando apenas o estado do Paraná, seus resultados do PIB foram superiores se comparados aos resultados do Brasil. Segundo IPARDES (2016), entre janeiro a junho de 2016, o retrocesso do PIB no Estado foi de apenas de 3%, sendo que no Brasil, foi de 4,6%. Considerando o período de junho de 2015 a junho de 2016, o recuo do PIB paranaense foi de 3,7%, mediante uma redução de 4,9% do

PIB brasileiro, revelando a necessidade de se estudar o comportamento do PIB regional para uma análise local mais específica. Logo, um estudo sobre a tendência do PIB numa mesoeconomia se torna essencial, mesmo considerando um curto período de tempo.

Neste sentido, esta pesquisa utilizou Análise Estatística Espacial de Dados de Área e métodos estatísticos de previsão, a partir de dados governamentais, para fornecer informações confiáveis pautada em metodologias econométricas a fim de apontar direções seguras de investimento futuro nas mais diversas áreas de atividades político-econômicas, abstendo-se de subjetividade e respaldando essas direções por estratégias estatísticas.

1.3 ESTRUTURA DO TRABALHO

Este estudo é composto por seis capítulos, incluindo a introdução. No Capítulo 2 é apresentada uma breve revisão de literatura, dando um curto relato sobre o PIB e seu deflator, da mesma forma, sobre as técnicas estatísticas utilizadas nesta pesquisa. No terceiro capítulo, apresenta-se a metodologia utilizada para definir as variáveis utilizadas e, também, as abordagens para a criação dos modelos de previsão. No Capítulo 4 tem-se a apresentação dos resultados. A conclusão do trabalho é apresentada no quinto capítulo e, finalmente, no capítulo 6, apresentam-se as considerações finais e recomendações.

2 REVISÃO DE LITERATURA

2.1 PRODUTO INTERNO BRUTO

O Produto Interno Bruto (PIB) é conhecido como principal indicador econômico utilizado para medir a atividade econômica de uma região em um período de tempo, sendo o montante, em valores monetários, dos usos finais de todos os bens e serviços produzidos na região em consideração, avaliados a preço de mercado, descontando as despesas com insumos empregados na fabricação dos produtos. Os setores responsáveis por esse cálculo são: a agropecuária, a indústria e o setor de serviços. Segundo IBGE (2015), o PIB pode ser calculado e expresso por três óticas diferentes, apresentando o mesmo resultado, sendo essas óticas:

- 1) do lado da produção - o PIB é igual ao valor da produção menos o consumo intermediário mais os impostos, líquidos de subsídios, sobre produtos não incluídos no valor da produção.
- 2) do lado da demanda - o PIB é igual a despesa de consumo final mais a formação bruta de capital fixo mais a variação de estoques mais as exportações de bens e serviços menos as importações de bens e serviços.
- 3) do lado da renda - o PIB é igual a remuneração dos empregados mais o total dos impostos, líquidos de subsídios, sobre a produção e a importação mais o rendimento misto bruto, mais o excedente operacional bruto.

Como o valor dos bens intermediários (insumos) está incluído no valor dos bens finais, estes não são considerados, pois resultaria em dupla contagem.

O cálculo do PIB foi criado por Simon Kuznets na década de 1930, devido à falta de informações econômicas nesta época, até mesmo sendo considerado o período da Grande Depressão nos Estados Unidos. Kuznets procurou compreender melhor a distribuição de riqueza na sociedade estadunidense e julgou ser adequado desagregar os volumes de produção (*output*) e insumos (*input*) de acordo com os diferentes setores da indústria. No ano de 1937, o PIB também foi introduzido na Grã-Bretanha, sendo remodelado pelo britânico Richard Stone, o qual tornou o PIB um indicador mais claro e concreto. O Fundo Monetário Internacional (FMI) foi o responsável por disseminar seus conceitos para a sociedade mundial, chegando ao Brasil em 1948 sob a incumbência do Instituto Brasileiro de Geografia e Estatística (IBGE). Desde esse período, o PIB se tornou o principal indicador econômico regional (COYLE, 2014; PHILIPSEN, 2015; MORENO, CHASSÉ; FUHR, 2016).

Entretanto, Michels, Oliveira e Wollenhaupt (2013) alertam que o PIB é apenas um indicador de crescimento econômico, não devendo ser visto como um indicador de desenvolvimento econômico. Apesar dos termos estarem relacionados, o crescimento econômico é um fator quantitativo e desenvolvimento econômico é um fator qualitativo. Apesar de ser frequentemente mencionado na mídia que a quantidade de produto gerado em um país é uma medida do desenvolvimento, esta estatística é apenas a medida de crescimento de uma nação. Tanto o PIB total quanto o PIB *per capita* (PIB total dividido pela população da região) compõem o conjunto de indicadores econômicos que evidenciam o nível de desenvolvimento regional, mas não abrangem completamente tal conceito. O desenvolvimento da localidade corresponde ao beneficiamento social resultante do crescimento. O desenvolvimento econômico agrega o crescimento econômico ao bem-estar da população, através da diminuição de suas desigualdades financeiras, na redução dos níveis de desemprego e da melhoria geral da qualidade de vida em características como nutrição, saúde, moradia, educação e transporte. (MICHELS; OLIVEIRA; WOLLENHAUPT, 2013)

Como já dizia Kuznets (1934) referente ao PIB como indicador de desenvolvimento:

o bem-estar econômico não pode ser mensurado adequadamente, a menos que a distribuição pessoal da renda seja conhecida. E nenhuma medida de renda se compromete a estimar o reverso da renda, ou seja, a intensidade e o esforço no lucro da renda. O bem-estar de uma nação pode, portanto, dificilmente ser inferido a partir de uma medida de renda nacional (Kuznets, 1934, p. 6-7).

Para classificar o crescimento da região estudada, analisa-se o período objetivo com o período anterior. Por exemplo, para analisar o crescimento de um país durante o ano de 2016, compara-se o PIB de 2016 do país em questão com o PIB de 2015. Se o PIB do período objetivo for maior que o período anterior, diz-se que a região teve um crescimento econômico, caso contrário, denomina-se o período como recessão (GUIMARÃES; LIRA, 2017a).

A nível nacional, o período de recessão econômica, que corresponde à redução da economia do país, caracteriza-se por dois trimestres consecutivos de retração do PIB. Isso pode ser devido a vários fatores, como o declínio da produtividade, o aumento da taxa de desemprego, a redução da renda familiar, o declínio dos investimentos, entre outras componentes do PIB.

Devido ao aumento de preços ao longo dos anos, Blanchard (2012) separa o PIB de duas maneiras: PIB nominal e PIB real.

2.1.1 PIB nominal

Segundo Blanchard (2012, p. 21), “o PIB nominal é a soma das quantidades de bens finais multiplicada por seus preços correntes”. Deste modo, com o passar do tempo, o crescimento nominal do PIB é observado devido à produção da maioria dos bens nesse período, e a maioria dos preços dos bens aumentar conjuntamente.

2.1.2 PIB real

O PIB real, por sua vez, é a soma da produção dos bens finais multiplicada pelos preços constantes. Desta forma, o aumento dos preços ao longo do tempo é subtraído do PIB. Logo, para analisar o crescimento de uma determinada região, o uso do PIB real se torna mais apropriado (BLANCHARD, 2012, p. 21).

Assim, para calcular o PIB real, faz-se necessário o uso de um deflator a fim de obter-se os valores a preços constantes.

2.2 DEFLATOR DO PIB

Segundo Bastos e Araújo (2017), “o deflator do PIB é uma média dos deflatores dos seus componentes, ponderada pela participação de cada componente no PIB”. Desta forma, a partir do PIB nominal (*PIBN*):

$$PIBN = P_A \times Q_A + P_X \times Q_X + P_M \times Q_M, \quad (2.2.1)$$

em que P_A , P_X e P_M são, respectivamente, os deflatores da absorção interna, das exportações e das importações, e Q_A , Q_X e Q_M são, respectivamente, as quantidades da absorção interna, das exportações e das importações. Fazendo a razão entre cada componente do PIB nominal pelos índices de preços respectivos, obtém-se o PIB real (*PIBR*):

$$PIBR = \frac{P_A \times Q_A}{P_A} + \frac{P_X \times Q_X}{P_X} + \frac{P_M \times Q_M}{P_M} = Q_A + Q_X - Q_M \quad (2.2.2)$$

Ao dividir o PIB nominal, equação (2.2.1), pelo PIB real, equação (2.2.2), é obtido o deflator do PIB (D_{PIB}):

$$D_{PIB} = \frac{PIBN}{PIBR} = \frac{P_A \times Q_A + P_X \times Q_X + P_M \times Q_M}{Q_A + Q_X - Q_M}, \quad (2.2.3)$$

Portanto, para comparar o PIB de diferentes períodos, é necessário ajustar esta variável para que possa ser comparada entre diferentes momentos de tempo. Este procedimento é chamado de deflação, uma vez que o índice de preços (deflator) é usado para remover o efeito da inflação.

O deflator do PIB é um indicador que mede a variação média dos preços de todos os bens finais produzidos na região, sendo o preço de cada um deles ponderado implicitamente pelo seu peso no ano em curso (GUIMARÃES; LIRA, 2017a).

2.3 ÍNDICES DE EMPREGO

Devido a limitação em obter dados do PIB mensal quando se restringe a nível de municípios brasileiros, dado ao fato que os dados municipais sobre o volume total de produção de bens e serviços terem periodicidade anual, buscou-se contornar esta complicação, utilizando outros indicadores, também importantes, que medissem a atividade econômica com a finalidade de estudar o crescimento econômico para indicar se a economia passa por um ciclo de expansão ou retração. Assim, foram utilizados indicadores de emprego, que indicam as variações na economia da região. Neste sentido, Saboia (2001) identifica três fontes de dados que se complementam. São elas: a Pesquisa Mensal de Emprego (PME/IBGE); a Relação Anual de Informações Sociais (RAIS/MTE); e o Cadastro Geral de Empregados e Desempregados (CAGED/MTE).

A Pesquisa Mensal de Emprego (PME), conduzida pelo IBGE, foi uma pesquisa domiciliar mensal que usava um esquema de amostragem de aproximadamente 40 mil domicílios a fim de analisar seis regiões metropolitanas do Brasil, as regiões capitais do Rio de Janeiro, São Paulo, Porto Alegre, Belo Horizonte, Recife e Salvador. O tema principal do estudo era o trabalho, abrangendo os setores formal e informal, contendo algumas características demográficas e educacionais, a fim de proporcionar uma melhor compreensão deste tema (RIBAS; SOARES, 2010). Contudo, esta

pesquisa foi substituída pela Pesquisa Nacional por Amostra de Domicílios Contínua, a partir de fevereiro de 2016.

A Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD contínua) é também uma pesquisa de amostra obtida da amostra principal de 211.344 domicílios particulares permanentes, a fim de obter informações contínuas sobre a inclusão da população no mercado de trabalho em relação às características demográficas e educacionais. A periodicidade da publicação de indicadores relacionados ao trabalho é trimestral e tem cobertura geográfica: Brasil, grandes regiões, divisões federais, 20 regiões metropolitanas, que contêm municípios de capitais (Manaus, Belém, Macapá, São Luís, Fortaleza, Natal, João Pessoa, Recife, Maceió, Aracaju, Salvador, Belo Horizonte, Vitória, Rio de Janeiro, São Paulo, Curitiba, Florianópolis, Porto Alegre, Vale do Rio Cuiabá, e Goiânia) e Região Integrada de Desenvolvimento da Grande Teresina (CARVALHO, 2016).

O Cadastro Geral de Empregados e Desempregados (CAGED) foi estabelecido pelo governo federal no ano de 1965, criando um registro permanente de admissões e demissões de funcionários de acordo com o regime de Consolidação de Leis do Trabalho (CLT). O CAGED consiste em uma fonte significativa de informações sobre o mercado de trabalho nacional e regional, apresentando periodicidade mensal. Além disso, o cadastro também fornece estatísticas geograficamente desagregadas no nível municipal sobre os resultados de admissão e demissão de funcionários celetistas.

Como, tanto o CAGED quanto a PNAD contínua, têm seus dados apresentados nos períodos mensal e trimestral, respectivamente, essas fontes de dados permitem uma avaliação mais rápida das flutuações do mercado de trabalho do que a RAIS. Entretanto, “a RAIS pode ser considerada um censo administrativo sobre o mercado de trabalho, na medida em que todas as organizações legais (privadas e públicas) são obrigadas a declará-la anualmente” (FERREIRA; MATOS, 2008).

Apesar da RAIS ter sua periodicidade anual, esta base de dados tem uma maior abrangência que a PNAD contínua e o CAGED, uma vez que cobre a nível nacional não apenas o movimento daqueles que foram admitidos e desligados de mês para mês, mas, também, todos os empregados do ano base de 01 de janeiro a 31 de dezembro de cada ano. Portanto, os tipos de vínculo de trabalho considerados pela RAIS são da forma: estatutários, celetistas, temporários e avulsos. Além disso, a

desagregação dos dados por municípios permite um estudo mais individualizado para cada região estudada.

Ramos (2003) conclui em seu trabalho que “uma mesma taxa de crescimento do PIB pode ter como corolário diferenças não-desprezíveis nos impactos sobre o mercado de trabalho, tanto em termos setoriais quanto tendo como parâmetro a qualidade do emprego”. Assim, fazer uma comparação entre os indicadores de emprego e o PIB é altamente relevante.

2.4 CORRELAÇÃO DE SPEARMAN

A análise de correlação tem como finalidade avaliar o grau de associação entre duas variáveis, geralmente denotadas por X e Y, ou seja, mede a “força” de relacionamento entre as variáveis X e Y. Para quantificar a relação entre duas variáveis quantitativas este trabalho utilizou o coeficiente de correlação de Spearman (r_s). Esse coeficiente é utilizado em substituição ao coeficiente de correlação de Pearson nos casos em que não há a normalidade dos dados e ainda em situações envolvendo poucos pares de dados, pois, ao contrário do coeficiente de correlação de Pearson, esta estatística não requer a hipótese de linearidade entre as variáveis consideradas e não exige variáveis quantitativas ou que elas tenham uma distribuição normal bivariada, requer somente que as variáveis estejam dispostas de forma ordinal, para que os elementos estudados (indivíduos ou objetos) formem duas séries ordenadas (LIRA, 2004).

O coeficiente de correlação de Spearman entre duas variáveis ordinais quantitativas, X e Y, é dado por:

$$r_s = 1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)} \quad (2.4.1)$$

onde:

r_s é o Coeficiente de Correlação de Spearman;

$x_i = X_i - \bar{X}$;

$y_i = Y_i - \bar{Y}$;

n é o número de observações.

Segundo Lira (2004), a interpretação do valor da correlação está sujeita aos objetivos de sua aplicação e as razões pelas quais esta estatística é calculada.

Entretanto, a autora cita, também, que o coeficiente de correlação pode ser interpretado qualitativamente da seguinte forma:

se $0,00 < |r_s| < 0,30$, existe fraca correlação entre as variáveis analisadas;

se $0,30 \leq |r_s| < 0,60$, existe moderada correlação entre as variáveis analisadas;

se $0,60 \leq |r_s| < 0,90$, existe forte correlação entre as variáveis analisadas;

se $0,90 \leq |r_s| < 1,00$, existe correlação muito forte entre as variáveis analisadas.

2.5 ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS (AEDE)

Segundo Perobelli, Ferreira e Faria (2006) a Análise Exploratória de Dados Espaciais (AEDE) é um conjunto de técnicas que tem por objetivo descrever e visualizar distribuições espaciais, identificar observações atípicas ou *outliers* espaciais, detectar padrões de associação espacial (*clusters* espaciais) e sugerir organizações espaciais ou outras formas de heterogeneidade espaciais.

Neste trabalho, utilizou-se, mais particularmente, a estatística de associação espacial local (LISA) para identificar se existe um padrão espacial na região estudada.

2.5.1 Índice Local De Associação Espacial (LISA)

Para medir o nível de associação da região em análise, pode ser utilizado o índice LISA (Local Indicators Spatial Association). O LISA é uma ferramenta de grande utilidade na AEDE. Essa técnica tem sido empregada em diversas áreas de estudo, quando se há interesse em analisar a dependência espacial, pois o LISA univariado se caracteriza por apresentar uma autocorrelação espacial da mesma informação em diferentes regiões no espaço.

Talvez o LISA mais comum na análise espacial seja o índice I local de Moran. Este índice foi criado por Anselin em 1995 para verificar a existência de autocorrelação espacial na amostra ou população estudada e, a partir disso, identificar *clusters* ou aglomerados que possibilitam identificar áreas que apresentam um padrão específico associado ao seu posicionamento geográfico (ROMERO, 2006). O índice quantifica a influência de cada partição com os seus respectivos vizinhos. A ideia do índice é, basicamente, comparar a característica de uma partição i com a média das características de seus vizinhos.

Desta forma, o índice I local de Moran é uma medida da correlação onde a correlação negativa indica dispersão de valores semelhantes, a correlação positiva

indica o agrupamento de valores similares (alta ou baixa) e a correlação zero, aleatoriedade espacial completa. Segundo Anselin (1995) O índice de Moran Local pode ser definido por:

$$I_i = \left(\frac{z_i}{\sum_i z_i^2 / n} \right) \sum_j w_{ij} z_j \quad (2.5.1)$$

Onde z_i e z_j são os desvios em torno da média; n é o número de observações e w_{ij} representa os elementos da matriz de pesos espaciais.

Ao fazer a comparação entre as partições, este indicador permite classificar a autocorrelação dos municípios nas seguintes formas: valores positivos e média dos vizinhos positivos (H-H); valores negativos e média dos vizinhos negativa (L-L); valores positivos e média dos vizinhos negativa (H-L) e valores negativos e média dos vizinhos positiva (L-H) (MOURA; LIRA, 2011; MOURA, LIRA; CINTRA, 2012).

2.6 ANÁLISE DE REGRESSÃO

O termo “Regressão” originou-se quando Francis Galton, em seu artigo “*Family likeness in statur*”, estudou o relacionamento entre as alturas de pais e filhos. Ele verificou que a média da estatura dos filhos de um grupo de pais altos era inferior a estatura de seus pais, e a estatura média de um grupo de filhos de pais baixos era superior do que a de seus pais; assim, filhos de pais altos e baixos “regrediam” do mesmo modo à altura média de todos os homens. Com esta conclusão, Galton disse que este acontecimento era uma “regressão à mediocridade” (GUJARATI; PORTER, 2011).

Na atualidade, a interpretação da regressão é bem diferente. Montgomery, Peck e Vining (2012) descrevem a regressão linear como uma técnica estatística com a finalidade de investigar e modelar a relação entre diferentes variáveis. De modo geral pode-se dizer que a análise de regressão é o estudo de uma variável (a variável dependente) em função de uma ou mais variáveis (as variáveis independentes), com o objetivo de estimar e/ou prever a média populacional ou valor médio da variável dependente, utilizando valores observados por amostragem das variáveis independentes (GUJARATI; PORTER, 2011).

Segundo Figueiredo Filho et al. (2011), atualmente a análise de regressão é uma das ferramentas ou métodos estatísticos utilizados com maior frequência na

ciência política contemporânea. Esta, é uma metodologia estatística para prever valores de uma variável resposta (dependente) para uma coleção de valores de variáveis regressoras (independentes), tendo por finalidade estimar e/ou prever o valor médio da primeira em relação aos valores conhecidos ou fixados (em amostragens repetidas) das segundas (GUJARATI; PORTER, 2011).

Montgomery, Peck e Vining (2012) descrevem a equação característica de um modelo de regressão linear múltipla com k variáveis regressoras, ou preditoras, que é dada pela equação (2.6.1)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (2.6.1)$$

sendo y a variável resposta, ou dependente; x_j ($j = 0, 1, \dots, k$) as variáveis regressoras, ou independentes; os parâmetros β_j ($j = 0, 1, \dots, k$) os coeficientes parciais de regressão, e ε o erro aleatório.

Segundo Charnet et al. (1999) os erros aleatórios constituem das inúmeros causas que, associadamente, podem afetar nas observações da variável dependente y .

O método mais habitual para estimar os coeficientes parciais de regressão é o método de mínimos quadrados. Este método consiste em encontrar uma estimativa para os parâmetros de forma que a soma do quadrado dos erros seja mínima. Os estimadores produzidos por este método são consistentes e não-viesados (FIGUEIREDO FILHO et al., 2011).

A ideia básica da Análise de Regressão é que, se o modelo for apropriado, os resíduos devem refletir as propriedades de seguirem uma distribuição normal, com variância constante (homocedasticidade) e média zero, serem independentes e os coeficientes serem estatisticamente significativos. Além disso, para regressão linear múltipla, deve ser feito o teste de coeficiente de determinação múltiplo a fim de conhecer a proporção da variação explicada da variável dependente pelas variáveis independentes.

Para testar se os resíduos seguem uma distribuição normal, existem vários testes estatísticos, tais como o teste de Shapiro-Wilk, Jarque-Bera, Kolmogorov-Smirnov, o gráfico Q-Q plot, entre outros já tão conhecidos na literatura. Tanto o teste de Shapiro-Wilk, Jarque-Bera e Kolmogorov-Smirnov tem como hipótese nula a normalidade dos dados, desta forma, se o p-valor for maior que o nível de significância

estipulado, não se rejeita a hipótese nula, caso contrário rejeita-se a hipótese de normalidade dos dados. (LEOTTI; COSTER; RIBOLDI, 2012, pg. 227-234).

Para a análise da homocedasticidade, Gujarati e Porter (2011) recomendam o uso do teste de Breusch-Pagan (BP). Este teste consiste em testar a hipótese nula de que as variâncias dos resíduos são iguais. Porém, a estatística BP é indicada apenas quando a distribuição dos resíduos é normal. (GUJARATI; PORTER, 2011, pg. 601-602).

O pressuposto de independência dos resíduos pode ser testado pela estatística de Durbin-Watson. O teste de Durbin-Watson tem como hipótese nula que a relação entre dois resíduos consecutivos é estatisticamente igual a zero (GUJARATI; PORTER, 2011, pg. 435-439).

Além dos testes para verificar a pressuposição do modelo estimado, deve-se testar a significância dos coeficientes estimados, a fim de verificar se cada uma das variáveis independentes individuais é significativa. Sendo a hipótese nula que o coeficiente avaliado é estatisticamente igual a zero. Usa-se o teste t para confirmar se a variável da regressão está realmente contribuindo para a estimativa (GUJARATI; PORTER, 2011, pg. 261).

2.7 ANÁLISE DE SÉRIES TEMPORAIS

Uma série temporal é um conjunto de observações dos valores que uma variável apresenta em períodos de tempo consecutivos em um intervalo estabelecido e que possuem uma dependência serial, isto é, dependência entre instantes de tempo. Esses dados podem ser coletados em intervalos de tempo uniformes (diariamente, semanalmente, mensalmente, trimestralmente, anualmente, etc.). Por vezes, os dados estão disponíveis em séries trimestrais e anuais, como no caso do PIB (GUJARATI; PORTER, 2011; STEFFEN, 2010).

Se os valores futuros de uma série temporal são, exatamente, determinados por alguma função matemática, a série temporal é dita determinista. Se os valores futuros podem ser descritos apenas em termos de uma distribuição de probabilidade, ou melhor, existe um componente aleatório independente, a série temporal é dita não determinista (ou estocástica).

Geralmente uma série temporal é representada por um conjunto de n observações $\{Z_1, Z_2, Z_3, \dots, Z_t, \dots, Z_{n-1}, Z_n\}$, em que, para cada t fixo, $t \in T$, sendo T um conjunto arbitrário, tem-se uma variável aleatória Z_t , que terá certa distribuição de probabilidades, este conjunto de observações, é chamado de família de variáveis aleatórias e é definido como processo estocástico. Na maioria das vezes, refere-se ao parâmetro t como sendo tempo, porém, a série Z_t poderá estar em função de algum outro parâmetro físico, como espaço ou volume (MORETIN; TOLOI, 2006).

Segundo Steffen (2010), a principal característica de uma série temporal, é que as observações vizinhas são dependentes e justamente esta dependência é que se procura analisar e modelar. Enquanto que em modelos de regressão, por exemplo, a ordem dos dados observados é irrelevante para a análise, em séries temporais a ordem das observações é de extrema importância.

Uma das suposições mais frequentes que se faz a respeito de uma série temporal é a de que ela é estacionária, ou seja, sua média e variância não variam sistematicamente ao longo do tempo. Entretanto, na prática, são poucas as séries que apresentam estacionariedade, em geral, são encontradas séries que apresentam tendências, sendo o caso mais simples as séries que apresentam tendência linear, isto é, a série flutua ao redor de uma reta que tem inclinação positiva ou negativa.

Uma série pode ser estacionária durante um longo período ou pode ser estacionária somente em períodos curtos, mudando de nível ou inclinação. Porém, a classe dos modelos ARIMA, que será descrito mais adiante, é capaz de descrever de maneira satisfatória séries estacionárias e séries não estacionárias, desde que não tenham comportamento explosivo (MORETIN; TOLOI, 2006)

Devido a maior parte dos métodos de análise estatística de séries temporais pressuporem que as séries sejam estacionárias, caso não sejam, é preciso transformar os dados originais de modo a tornar os dados estacionários. Moretin e Toloi (2006) citam que o caso mais habitual é tomar diferenças consecutivas da série original, sendo que a primeira diferença é feita da seguinte maneira:

$$\nabla Z_t = Z_t - Z_{t-1}. \quad (2.7.1)$$

Considerando um operador de translação para o passado, denotado por B e definido por

$$BZ_t = Z_{t-1}, B^m Z_t = Z_{t-m}, \quad (2.7.2)$$

tem-se que a primeira diferença corresponde a

$$\nabla Z_t = (1 - B)Z_t. \quad (2.7.3)$$

De modo geral, a d-ésima diferença de Z_t é:

$$\nabla^d Z_t = (1 - B)^d Z_t. \quad (2.7.4)$$

Normalmente, para a série ficar estacionária é suficiente diferenciar uma ou duas vezes.

Os modelos de previsão desenvolvidos por Box & Jenkins, foram elaborados na década de 1970 baseados no trabalho de Yule, que desenvolveu modelos de autorregressão (AR) em 1926 e modelos de média móvel (MA), adotados por Slutsk em 1937. A partir daí, Box & Jenkins construíram um método para identificar o modelo, estimar os parâmetros e validar o modelo. Devido a associação dos modelos AR e MA, os modelos Box & Jenkins são também chamados de modelos autorregressivos integrados de médias móveis (ARIMA) (STEFFEN, 2010).

Os modelos estocásticos empregados neste trabalho, baseiam-se na ideia de que uma série temporal Z_t em que valores sucessivos são dependentes, pode frequentemente ser considerada como gerada a partir de uma série de "choques" independentes a_t . Esses choques aleatórios tem uma distribuição de probabilidade fixa, geralmente assumida normal, com média zero e variância σ_a^2 . Chama-se essa sequência de variáveis aleatórias independentes $a_t, a_{t-1}, a_{t-2}, \dots$ de ruído branco (BOX, JENKINS; REINSEL, 2008).

Segundo Moretin e Tolo (2006), os modelos ARIMA supõem que a série temporal seja gerada por um filtro linear ou sistema linear, cujo ruído branco a_t é supostamente transformado nos dados da série temporal Z_t . A operação de filtragem linear forma uma soma ponderada de choques aleatórios anteriores a_t , de modo que

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \psi(B)a_t, \quad (2.7.5)$$

em que

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots \quad (2.7.6)$$

é denominada de função transferência do filtro, e μ é o parâmetro que determina o nível da série.

A sequência formada pelos pesos ψ_i pode ser finita ou infinita. Se esta sequência é finita, ou infinita e absolutamente convergente, o filtro é dito estável e o processo Z_t é estacionário e o parâmetro μ é a média sobre a qual os dados variam. Caso contrário, a série Z_t não é estacionária e μ não possui um significado específico, a não ser como ponto de referência para o nível do processo (BOX; JENKINS; REINSEL, 2008).

Como dito anteriormente, os modelos ARIMA são modelos compostos por modelos autorregressivos (AR), integrados e com modelos de médias móveis (MA). Os modelos autorregressivos são um caso particular de um filtro linear. A principal característica do modelo autorregressivo é que, em vez de variáveis independentes, o processo será usado para os valores anteriores da série para estimar o modelo (LIEBEL, 2004). Neste modelo, o valor atual da série é expresso como um associado finito e linear de valores anteriores do processo e um choque aleatório a_t , como mostra a equação (2.7.7):

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t. \quad (2.7.7)$$

A equação (2.7.7) é chamada de modelo autorregressivo de ordem p e denotado por $AR(p)$, onde:

$$\tilde{Z}_i = Z_i - \mu;$$

ϕ_i corresponde ao i -ésimo parâmetro autorregressivo, $i = 1, 2, \dots, p$;

a_t é o erro (ruído) no instante t .

Quando se quer definir um operador autorregressivo de ordem p em termos do operador de translação para o passado B , faz-se:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \quad (2.7.8)$$

Então o modelo autorregressivo descrito na equação (2.7.7) pode ser escrito como

$$\phi(B)\tilde{Z}_t = a_t. \quad (2.7.9)$$

Além da parte AR(p) dos modelos ARIMA, existe a parte composta por médias móveis (MA), denotada por MA(q), a qual consistem em considerar o processo linear da equação (2.7.5) supondo que $\psi_i = 0, j > q$. A notação dos pesos do filtro linear de médias móveis é dada por θ_i , sendo o parâmetro a ser estimado (MORETIN; TOLOI, 2006). Deste modo, tem-se:

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.7.10)$$

e, sendo $\tilde{Z}_t = Z_t - \mu$, o modelo é descrito como

$$\tilde{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t, \quad (2.7.11)$$

onde

$$\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q) \quad (2.7.12)$$

é o operador de médias móveis de ordem q .

Existem ainda os modelos mistos autorregressivos e de médias móveis ARMA(p, q), que podem ser escritos de maneira simplificada utilizando os operadores autorregressivo e de médias móveis, como mostra a equação (2.7.13) abaixo:

$$\phi(B)\tilde{Z}_t = \theta(B)a_t. \quad (2.7.13)$$

Entretanto, os modelos AR(p), MA(q) e ARMA(p,q) supõem a estacionariedade dos dados para ajustar o modelo. Liebel (2004) afirma que a não estacionariedade não impede que uma série não apresente algum tipo de homogeneidade. Este comportamento homogêneo pode ser representado pelo operador autorregressivo genérico $\varphi(B)$ de ordem d , que é dado por:

$$\varphi(B) = \phi(B)(1 - B)^d \quad (2.7.14)$$

onde $\phi(B)$ é um operador estacionário autorregressivo. Assim, um modelo que pode representar comportamento não estacionário homogêneo é da forma

$$\varphi(B)Z_t = \phi(B)(1 - B)^d Z_t = \theta(B)a_t \quad (2.7.15)$$

ou seja,

$$\varphi(B)\omega_t = \theta(B)a_t \quad (2.7.16)$$

onde

$$\omega_t = (1 - B)^d Z_t = \nabla^d Z_t. \quad (2.7.17)$$

Assim, algumas vezes o comportamento não estacionário homogêneo pode ser representado por um modelo que exige que a d-ésima diferença do processo seja estacionária (BOX; JENKINS; REINSEL, 2008).

Se $\omega_t = \nabla^d Z_t$ for estacionária, podemos representar ω_t por um modelo ARMA (p,q). Se ω_t for uma diferença de Z_t , então Z_t é uma integral de ω_t , por isso diz-se que Z_t segue um modelo autorregressivo, integrado e de médias móveis (ARIMA) de ordem (p,d,q) (MORETIN; TOLOI, 2006).

Em muitos casos, a série temporal apresenta sazonalidade, devendo-se acrescentar essa informação ao modelo, que será dada por

$$\Phi(B^s)\nabla_s^D Z_t = \theta(B^s)\alpha_t \quad (2.7.18)$$

onde:

s é o período de sazonalidade;

$\Phi(B^s)$ é o operador autorregressivo sazonal de ordem P, estacionário;

$\theta(B^s)$ é o operador sazonal de médias móveis de ordem Q, invertível;

$\nabla_s^D = (1 - B^s)^D$ é o operador de diferença sazonal e D indica o número de diferenças sazonais;

α_t é o componente de erro.

Adicionando a informação de sazonalidade ao modelo ARIMA¹, um modelo autorregressivo integrado à média móvel sazonal (SARIMA) multiplicativo de ordem (p,d,q) X (P,D,Q)_s, é dado por

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D Z_t = \theta(B)\theta(B^s)\alpha_t \quad (2.7.19)$$

Gujarati e Porter (2011) citam que a metodologia Box & Jenkins consiste em quatro etapas, sendo elas:

¹ Detalhes podem ser encontrados em Moretin e Toloi (2006, p. 249-251)

1. Identificação do modelo – o objetivo é determinar os valores de p , d e q do modelo $ARIMA(p,d,q)$;
2. estimação dos parâmetros do modelo;
3. verificação do diagnóstico – verifica-se se os resíduos são ruídos brancos, se a resposta for positiva pode-se passar para a última fase, caso contrário, deve-se voltar a primeira etapa;
4. previsão – depois de identificado o modelo, estimado e verificado, pode-se realizar previsões para períodos futuros.

Gujarati e Porter (2011) afirmam que as principais técnicas na fase da identificação são a função de correlação amostral (ACF), a função de correlação amostral parcial (PACF) e os seus correlogramas, que representam a ACF e a PACF contra a extensão de defasagem. Steffen (2010, p. 41) cita que “a análise gráfica da FAC e FACP é uma estratégia muito utilizada pelos analistas de séries temporais para identificação dos modelos”. Os valores de ACF e PACF de uma série temporal, bem como suas representações gráficas, são facilmente calculados por *softwares* estatísticos, tais como *Statgraphics*, *Minitab*, e *R development core team*.

A etapa de estimação dos parâmetros utiliza-se do algoritmo de Levenberg-Marquardt que calcula estimativas de mínimos quadrados não-lineares de forma otimizada matematicamente. Este é um algoritmo iterativo, no qual é inviável calcular sem o auxílio computacional.

A terceira etapa da metodologia de Box & Jenkins é a etapa de diagnóstico do modelo estimado, verificando através dos resíduos se o modelo corresponde aos dados da série temporal ou um modelo alternativo deve ser buscado. Geralmente esses testes são baseados nas autocorrelações estimadas dos resíduos (MORETIN; TOLOI, 2006). Os objetivos desses testes é que os resíduos do modelo ajustados sejam ruídos branco. Box, Jenkins e Reinsel (2008) indicam os testes da autocorrelação dos resíduos, *Portmanteu tests*, normalidade dos resíduos e periodograma acumulado.

O teste de autocorrelação residual (ACF dos resíduos) é um teste utilizado para avaliar se há correlação entre os resíduos. Quando o modelo estimado é adequado, os resíduos devem ser aproximadamente não correlacionados. Desta

forma, o gráfico da ACF para os resíduos deve mostrar valores dentro do intervalo dentro do qual a autocorrelação é considerada nula.

Segundo Minervi (2013) os *Portmanteu tests* são uma família de testes estatísticos, dos quais se destacam o teste de Box e Pierce e o teste de Ljung e Box. Em 1970, Box e Pierce “sugeriram um teste para as autocorrelações dos resíduos, que apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos” (MORETIN; TOLOI, 2006). Este teste considera o conjunto dos k primeiros coeficientes de autocorrelação (r_k) de uma só vez. O resultado da estatística de Box-Pierce (Q) é comparado com os valores da distribuição Qui-Quadrado (χ^2) para verificar se a estatística é significativamente diferente de zero. Entretanto, em 1978, uma variação desse teste foi sugerida por Ljung e Box.

O teste de Ljung-Box afirma que, se o modelo for adequado, a estatística

$$\tilde{Q}(k) = n(n+2) \sum_{i=1}^k \frac{r_i^2}{n-i} \quad (2.7.20)$$

terá aproximadamente uma distribuição χ^2 com $k - p - q$ graus de liberdade. A hipótese nula é a de que os resíduos são ruído branco (MINERVI, 2013).

Apesar de existirem vários testes para avaliar a normalidade dos resíduos, geralmente as pesquisas econométricas se utilizam do teste de Jarque-Bera (GUJARATI; PORTER, 2011).

O teste de normalidade de Jarque-Bera tem como hipótese nula de que os resíduos são normalmente distribuídos. Segundo Gujarati e Porter (2011), Jarque e Bera demonstraram que, em amostras grandes, a estatística JB dada na EQUAÇÃO (2.7.21), abaixo, segue a distribuição de qui-quadrado com 2 graus de liberdade (χ_2^2).

$$JB = n \left[\frac{S^2}{6} + \frac{(K-3)^2}{24} \right] \quad (2.7.21)$$

onde:

n é o tamanho da amostra;

S é o coeficiente de assimetria da distribuição, dado pela EQUAÇÃO (2.7.22);

K é o coeficiente de curtose² dado pela EQUAÇÃO (2.7.23).

$$S = \frac{1}{n} \sum \left[\frac{X_i - \bar{X}}{\hat{\sigma}} \right]^3 \quad (2.7.22)$$

$$K = \frac{1}{n} \sum \left[\frac{X_i - \bar{X}}{\hat{\sigma}} \right]^4 - 3 \quad (2.7.23)$$

Se o p-valor calculado para a estatística JB for menor que o nível de significância, pode-se rejeitar a hipótese de que a distribuição dos resíduos é normal. Entretanto, se o p-valor for maior que o nível de significância estipulado, não se rejeita a hipótese nula, de forma que a normalidade dos resíduos é considerada verdadeira.

Outra maneira de se diagnosticar os resíduos é a utilização do Teste do Periodograma Acumulado. Este teste foi projetado especificamente para detectar padrões periódicos no caso de ruído branco. Nesse teste compara-se o periodograma acumulado dos resíduos estimados com um ruído branco.

O periodograma da série temporal a_t , $t = 1, 2, \dots, n$ é dado por

$$I(f_i) = \frac{2}{n} \left[\left(\sum_{t=1}^n a_t \cos(2\pi f_i t) \right)^2 + \left(\sum_{t=1}^n a_t \sin(2\pi f_i t) \right)^2 \right] \quad (2.7.24)$$

onde $f_i = i/n$ é a frequência. Desta forma, o periodograma acumulado é uma forma de correlacionar os valores a_t com ondas de seno e cosseno de diferentes frequências. O periodograma acumulado também fornece meios efetivos para a detecção de períodos não aleatórios (BOX; JENKINS; REINSEL, 2008). Basicamente, a análise do periodograma acumulado é feita visualmente, já que os *softwares* estatísticos para séries temporais geralmente têm este teste implementado. Assim, basta verificar se o gráfico do periodograma acumulado dos resíduos se aproxima de uma reta, não contendo componentes periódicos. Caso não se aproxime da reta, significa que ainda existem partes sistemáticas a serem modeladas, devendo voltar à primeira etapa. Caso haja componentes periódicos, significa que os dados apresentam sazonalidade, devendo ser aplicados modelos SARIMA (p,d,q) x (P,D,Q).

² Curtose é uma medida de dispersão que caracteriza o "achatamento" da curva da função de distribuição.

Se o modelo ajustado foi diagnosticado como sendo um bom modelo, Bueno comenta que (2011, p.91), “a previsão é relativamente fácil e direta”. Moretin e Tolo (2006), comentam que é possível expressar a previsão \hat{Z}_{t+h} de três maneiras diferentes utilizando o modelo ARIMA, a previsão utilizando a equação de diferenças, a previsão utilizando a forma de choques aleatórios e a previsão utilizando a forma invertida. De acordo com a previsão utilizando a equação de diferenças, a equação de previsão, considerada como uma função de h , com origem em t fixo, deve satisfazer a seguinte equação

$$\hat{Z}_t(h) = \sum_{i=1}^{p+d} \phi \hat{Z}_t(h-1), \quad h > q. \quad (2.7.25)$$

Entretanto, deve-se considerar o fato de que a variância aumenta quanto mais se aumenta o horizonte de previsão, desta forma os intervalos de confiança para a previsão tendem a crescer (MINERVI, 2013).

Outra forma de modelar uma série temporal consiste em calcular a média aritmética das k observações mais recentes, esta abordagem é chamada de Médias Móveis Simples (MMS(k)). Na prática, escolhe-se uma quantidade k de períodos anteriores e calcula-se a média dos valores. Assim, a média móvel simples pode facilmente ser obtida através da EQUAÇÃO (2.7.26):

$$M_t = \frac{Z_t + Z_{t-1} + \dots + Z_{t-k+1}}{k} \quad (2.7.26)$$

Desta forma a cada período a observação mais antiga é substituída pela mais recente, calculando-se uma média nova, por esse motivo é denominado média móvel. Assim, a previsão dos h valores futuros é dada pela última média móvel calculada, sendo o melhor tamanho de k aquele que apresenta o menor erro quadrático médio de previsão (SILVA, 2008).

3 MATERIAL E MÉTODOS

3.1 MATERIAL

Os dados foram extraídos dos portais do Instituto Brasileiro de Geografia e Estatística (IBGE), do Ministério do Trabalho e Emprego (MTE) e, ainda, do site do Instituto Paranaense de Desenvolvimento Econômico e Social (IPARDES) desagregados em nível municipal, conforme os objetivos deste trabalho.

Do portal do IBGE, foram obtidos o arquivo *shapefile* e os dados do PIB a preços correntes por município. O arquivo *shapefile* foi usado no *software* GeoDa para a confecção dos mapas dos índices de Moran local, para identificação de *clusters* a partir dos quais foram selecionados os municípios analisados neste trabalho. Para os dados do PIB por municípios, foi elaborada uma planilha de 15 colunas por 400 linhas. A primeira linha contém os nomes das variáveis. A primeira coluna contém o código de cada município, a segunda coluna contém o nome de cada município paranaense e da terceira a décima quinta coluna contém os valores do PIB nominal municipal por ano, desde 2002 a 2014. Foi definido este período em virtude dos dados do PIB por municípios serem disponibilizados apenas neste período. Desse modo, a partir da segunda linha, tem-se os dados do PIB nominal municipal, totalizando 399 municípios.

Através do portal do MTE, adquiriu-se os microdados do CAGED e da RAIS, para a obtenção dos dados desagregados de emprego por municípios classificados como emprego celetista e estatutário. Primeiramente, foi criada uma planilha com o emprego total da RAIS e depois foi confeccionada uma planilha com apenas os dados de empregados celetistas de 15 municípios selecionados do estado do Paraná, através dos *clusters* formados pelo Índice de Moran local. Esta última com apenas 15 linhas e 16 colunas, na qual a primeira linha contém os nomes das variáveis e a primeira coluna refere-se ao ano (2002 a 2015) e as demais colunas são os dados do PIB por municípios.

Estes dados foram obtidos através da tabulação dos microdados da RAIS, considerando os empregados classificados como trabalhador: urbano vinculado a empregador pessoa jurídica por prazo indeterminado (código 10) , urbano vinculado a empregador pessoa física por prazo indeterminado (código 15) , rural vinculado a empregador pessoa jurídica por prazo indeterminado (código 20) , rural vinculado a empregador pessoa física por prazo indeterminado (código 25), urbano vinculado a

empregador pessoa jurídica por tempo determinado ou obra certa (código 60), urbano vinculado a empregador pessoa física por tempo determinado ou obra certa (código 65), rural vinculado a empregador pessoa jurídica por tempo determinado ou obra certa (código 70) e rural vinculado a empregador pessoa física por tempo determinado ou obra certa (código 75).

Através do site do IPARDES, foram obtidos os dados do consumo de Energia Elétrica e rendimento médio mensal dos trabalhadores paranaenses. Os dados das duas variáveis foram alocados em planilhas separadas, produzidas pelo próprio site do IPARDES, onde cada planilha é composta de 15 colunas e 400 linhas. A primeira linha contém os nomes das variáveis. A primeira coluna é composta pelos nomes dos municípios e as demais colunas compostas pelos dados anuais das respectivas variáveis por municípios. Desta forma, na linha, tem-se os dados anuais da variável referente ao consumo de energia elétrica e rendimento médio, por município.

3.2 MÉTODOS

Este estudo foi subdividido em três estágios, I, II e III. O estágio I, caracteriza-se pelo tratamento e seleção de dados, o estágio II pela Análise de Regressão Simples e Múltipla, e o estágio III pela Análise de Séries Temporais.

3.2.1 Estágio I – Tratamento e seleção de dados

Devido os dados obtidos do PIB serem em valores nominais, buscou-se remover a inflação acometida ao longo do tempo utilizando um deflator concedido pelo IPARDES com ano base em 2014, para se obter o PIB municipal real para os anos estudados. Desta forma, foi confeccionada uma nova planilha de 15 colunas por 400 linhas, com a primeira linha contendo os nomes das variáveis, a primeira e a segunda colunas contendo o código e o nome de cada município paranaense, respectivamente, e da terceira a décima quinta coluna contendo os valores do PIB municipal a preços constantes (2014 = 100) por ano, desde 2002 a 2014.

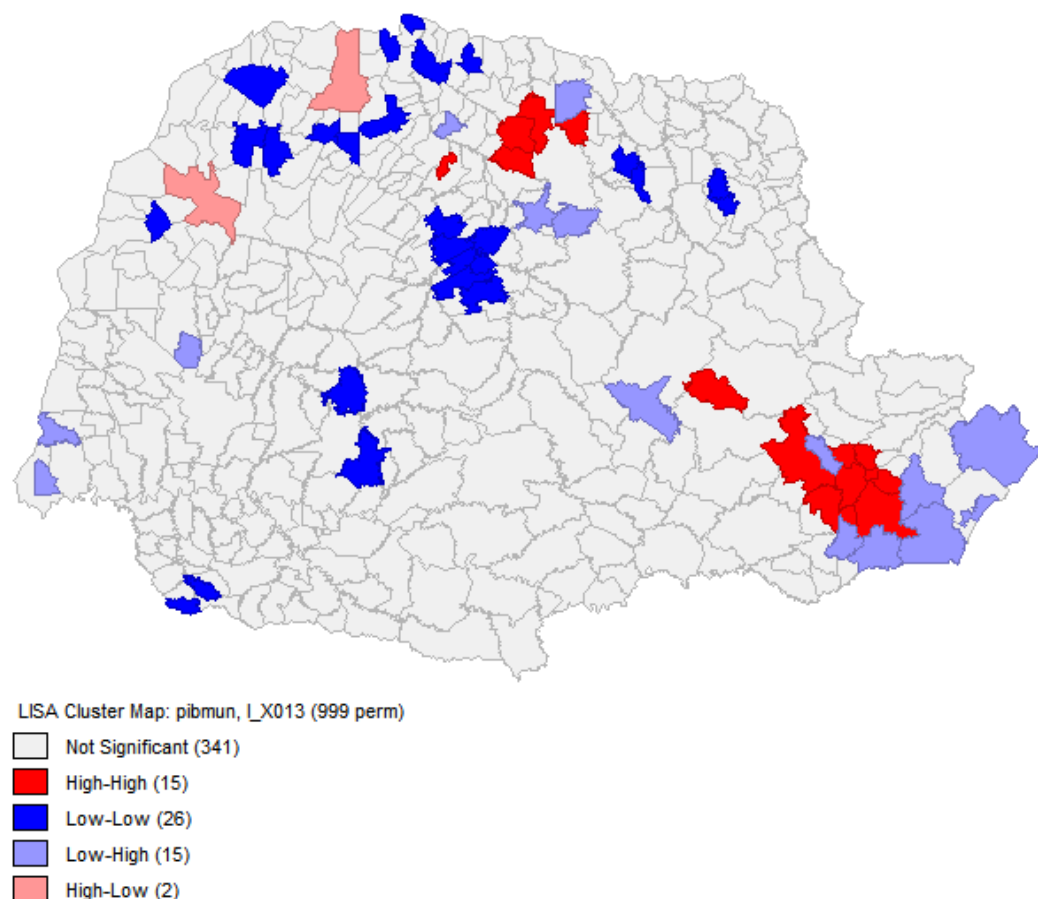
No contexto do Paraná, são 399 municípios divididos geograficamente em 10 regiões: Centro-ocidental paranaense, Centro-oriental paranaense, Centro-sul Paranaense, Metropolitana de Curitiba, Noroeste paranaense, Norte central paranaense, Norte pioneiro paranaense, Oeste paranaense, Sudeste paranaense e Sudoeste paranaense. Conhecer regiões com características semelhantes e também

aquelas com diferenças, é primordial para a gestão pública. Essas semelhanças, ou as diferenças, podem ser identificadas através de técnicas estatísticas espaciais de área.

Com os dados do PIB real dos 399 municípios paranaenses, verificou-se que o ano de 2013 foi o ano que apresentou melhor índice de crescimento a preços constantes, então partiu-se para uma Análise Exploratória de Dados Espaciais (AEDE) utilizando o Índice Local de Associação Espacial (LISA) para selecionar alguns municípios a fim de aplicar a proposta deste trabalho.

A AEDE foi feita utilizando o *software* livre GeoDa. Primeiramente, tomou-se o arquivo *shapefile* para identificar os municípios do estado do Paraná, e em seguida usou-se os dados do PIB real de 2013 para formação de *clusters* utilizando o LISA pelo GeoDa. Através deste indicador os municípios foram classificados em quatro categorias de correlação diferentes, sendo elas: alta-alta (H-H), baixa-baixa (L-L), alta-baixa (H-L) e baixa-alta (L-H), como mostra a FIGURA 1 a seguir:

FIGURA 1 – MAPA DE CLUSTER LISA DO PIB A PREÇOS CONSTANTES (2014 = 100) DOS MUNICÍPIOS PARANAENSES – 2013



FONTE: O autor (2018).

A partir destas classificações, foram escolhidos *clusters* de municípios que apresentaram valores do LISA positivos e média dos valores dos municípios vizinhos também positivos. Alguns exemplos de classificação H-H de municípios são apresentados na TABELA 1 a seguir:

TABELA 1 – MUNICÍPIOS LIMÍTROFES, INDICADOR PADRONIZADO, MÉDIA DO INDICADOR PADRONIZADO DOS MUNICÍPIOS LIMÍTROFES E CLASSIFICAÇÃO, SEGUNDO ALGUNS MUNICÍPIOS, PARANÁ – 2013

MUNICÍPIO	MUNICÍPIOS LIMÍTROFES	INDICADOR PADRONIZADO DO MUNICÍPIO (Z)	MÉDIA DO INDICADOR PADRONIZADO DOS MUNICÍPIOS LIMÍTROFES (Wz)	CLASSIFICAÇÃO
Curitiba		17,9053	1,1427	H-H
	Almirante Tamandaré	0,0602		
	Araucária	1,3616		
	Campo Largo	0,6003		
	Campo Magro	-0,1266		
	Colombo	0,6706		
	Fazenda Rio Grande	0,1111		
	Pinhais	0,9356		
	São José dos Pinhais	5,5286		
	Média dos Municípios limítrofes	1,1427		
Araucária		1,3616	2,6018	H-H
	Balsa Nova	-0,0645		
	Campo Largo	0,6003		
	Contenda	-0,1343		
	Curitiba	17,9053		
	Fazenda Rio Grande	0,1111		
	Mandirituba	-0,0749		
	Quitandinha	-0,1306		
	Média dos Municípios limítrofes	2,6018		
Arapongas		0,6781	0,8904	H-H
	Apucarana	0,3800		
	Cambé	0,5023		
	Londrina	3,4336		
	Rolândia	0,2556		
	Sabáudia	-0,1196		
	Média dos Municípios limítrofes	0,8904		

FONTE: O autor (2018).

Usando este método de seleção, foram selecionados 15 municípios com classificação H-H, no estado do Paraná, sendo eles: Almirante Tamandaré,

Arapongas, Araucária, Cambé, Campo Largo, Carambeí, Colombo, Curitiba, Fazenda Rio Grande, Ibiporã, Pinhais, Piraquara, Rolândia, São José dos Pinhais e Sarandi.

Para selecionar a variável do banco de dados mais correlacionada com o PIB, utilizou-se o *software* estatístico *R development core team* para calcular a correlação de Spearman (r_s) entre rendimento médio mensal dos trabalhadores paranaenses, consumo de energia elétrica total, industrial e comercial, emprego celetista e o PIB real de 2013, sendo a escolha deste ano também pelo motivo de ter sido o ano que apresentou maior PIB a preços constantes (2014=100) do estado do Paraná, no período analisado.

Entre as variáveis, foi escolhida a variável emprego celetista da RAIS, por ter apresentado a maior correlação de Spearman ($r_s = 0,9286$), sendo que os resultados da correlação de Spearman com as variáveis é apresentada no APÊNDICE A. Provavelmente esta variável obteve maior correlação com o PIB pois os empregados regidos pela Consolidação das Leis do Trabalho (CLT) é o grupo que está mais suscetível à altas e baixas do mercado de trabalho pois o empregado não obtém direito a estabilidade como no Regime Estatutário (GUIMARÃES; LIRA, 2017b).

Entretanto, para uma obtenção de uma série temporal maior com base nos dados da RAIS, foi elaborada uma nova planilha de dados, com 97 linhas, sendo a primeira o nome de cada município, e 15 colunas, as quais obtém os dados por município de forma que os valores de emprego celetista da RAIS são os parâmetros base para o final do mês de dezembro, e nos outros meses foi utilizado os dados do CAGED para ir somando os valores da RAIS mês a mês. Contudo, os dados do CAGED têm uma pequena diferença com os dados da RAIS, devido a isso, a proposta foi de somar a cada mês, a diferença proporcional entre o valor da RAIS aos meses já somados pelo CAGED.

Por exemplo, segundo a RAIS, no dia 31 de dezembro do ano de 2006 a cidade de Almirante Tamandaré tinha 7.244 pessoas empregadas pelo regime da CLT, e no dia 31 de dezembro de 2007, o mesmo município tinha 8.147 empregados celetistas. Um aumento de 903 empregados de um ano para outro. Pelo CAGED, o aumento foi de 849 empregados nesse mesmo período. Logo, o CAGED não registrou 54 admissões, se comparado à RAIS. Então estes 54 empregados foram igualmente distribuídos ao longo do ano. Obtendo, assim, dados mensais com referência anual

na RAIS e com acréscimos mensais do CAGED mais a razão da diferença anual entre RAIS e CAGED pelo total de meses no ano.

3.2.2 Estágio II – Análise de Regressão Simples e Múltipla

Após a definição dos municípios a serem estudados e da variável a ser relacionada com o PIB, recorreu-se à técnica de regressão linear para estimar uma relação funcional entre o emprego celetista e o PIB.

Os modelos foram decididos com o auxílio do *software* livre *R development core team* a partir dos dados por município definidos no Estágio I desta pesquisa, utilizando o período de 2002 a 2014 tanto para a variável resposta PIB real quanto para a variável explicativa emprego celetista obtido através da RAIS. Devido às diferentes características dos municípios, foram ajustados diferentes modelos, sendo eles: regressão linear simples; exponencial; e regressão linear múltipla, no último caso, fazendo uso de variável *dummy*³ para identificar os anos em que os fatores econômicos regionais foram atípicos.

Apresenta-se, como exemplo, na TABELA 2, as variáveis envolvidas no ajuste do modelo de regressão para o município de Araucária. Os demais municípios foram tratados de forma semelhante, sendo que, em situações onde não foram utilizadas a variável *dummy*, tem-se somente as variáveis ano, PIB e emprego celetista.

Com base nisso, para o ajuste da relação entre PIB e emprego celetista, assumiu-se três modelos que compreendem a regressão linear simples (tendo como variável independente o emprego celetista), regressão linear múltipla (tendo como variáveis independentes o emprego celetista e *dummy*) e exponencial linearizada:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 RAIS_CLT + \hat{\beta}_2 Dummy, \quad (3.2.2.1)$$

e

$$\ln(\hat{y}) = \hat{\beta}_0 + \hat{\beta}_1 RAIS_CLT \quad (3.2.2.2)$$

³ *Dummy* é uma variável qualitativa, frequentemente denominada variável binária ou dicotômica, visto que assume somente dois valores (0 ou 1) para indicar a presença ou ausência de uma qualidade ou atributo (GUJARATI; PORTER, 2011).

onde: \hat{y} representa a variável dependente (PIB a preços constantes estimado), $RAIS_CLT$ e $Dummy$ são variáveis explicativas, sendo emprego celetista obtido pela RAIS e $dummy$, respectivamente e $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ os coeficientes estimados.

TABELA 2 – DADOS DO PIB, EMPREGO CELETISTA E DUMMY PARA O MUNICÍPIO DE ARAUCÁRIA – 2002-2014

ANO	PIB A PREÇOS CONSTANTES (R\$ 1.000)	EMPREGO FORMAL	DUMMY
2002	7.654.020	19.646	1
2003	8.522.567	20.516	1
2004	9.660.964	22.367	1
2005	9.435.647	23.103	1
2006	10.089.252	25.478	1
2007	10.381.626	27.307	1
2008	11.369.757	32.108	1
2009	10.954.705	38.658	1
2010	12.793.666	46.669	1
2011	9.565.293	46.087	0
2012	6.710.536	39.387	0
2013	7.247.157	38.047	0
2014	8.558.227	37.162	0

FONTE: IBGE, RAIS.

Com os modelos de regressão ajustados, foi feita a análise de resíduos para testar se os modelos encontrados são válidos. Como mencionado na revisão de literatura, a análise sugere que, se os modelos forem adequados, os resíduos devem seguir uma distribuição normal, ter média nula, variância constante (homocedasticidade) e serem independentes. Então, foram utilizados os testes de: Shapiro-Wilk, para testar a normalidade dos resíduos; Breush-Pagan, para a verificação da homocedasticidade e, ainda, o teste de Durbin-Watson para a verificação de independência.

Para que os testes fossem considerados válidos, foi utilizado o nível de significância de 5%, ou seja, o p-valor para os testes devem ser maiores que 0,05 para que os testes fossem aceitos.

Após os modelos terem sido ajustados e validados, foram todos armazenados em um arquivo *Rscript* a fim de serem utilizados para a projeção do PIB do ano de 2015, após serem encontradas as projeções do emprego celetista no estágio III.

3.2.3 Estágio III – Análise de Séries Temporais

Com a variável emprego celetista da planilha de relação entre RAIS e CAGED proposta no primeiro estágio da pesquisa, utilizou-se a metodologia de Box & Jenkins, com sazonalidade, usando os *softwares Statgraphics Centurion XVI* e o *R development core team* para fazer a análise de séries temporais.

Os dois *softwares* utilizados têm a metodologia de previsão automática dos modelos SARIMA(p,d,q)X(P,D,Q) implementadas em suas bibliotecas de previsão. Entretanto, ao serem propostos os modelos pelos softwares, na etapa de diagnóstico dos modelos, a maioria deles foi rejeitada devido a não ser validado em algum dos testes para aceitação do modelo ajustado. Por este motivo, viu-se a necessidade de programar uma nova função que além de propor o modelo, fizesse todos os testes de validação do mesmo.

A função foi escrita em *Rscript* e denominada *best.arima()*⁴, em que os dados de entrada são os dados da série temporal que se deseja ajustar o modelo e a ordem máxima dos parâmetros que se deseja ajustar o modelo. O máximo estipulado para ajuste de modelo foi SARIMA(5,2,5)X(3,2,3).

A função *best.arima()* retorna duas possíveis respostas para os dados de saída. A primeira resposta é “NÃO FOI ENCONTRADO MODELO ALGUM” para quando a função não encontrar nenhum modelo que seja validado em todos os testes descritos na seção 6.2. A segunda resposta é “O melhor modelo é ARIMA(p, d, q)X(P, D, Q)”, pois, entre vários modelos testados e aprovados, a função apresenta o modelo com menor critério AIC⁵, juntamente com os gráficos da ACF dos resíduos e o gráfico do periodograma acumulado.

Mesmo com a função *best.arima()*, para os municípios de Araucária, Fazenda Rio Grande e Pinhais, não foram encontrados modelos SARIMA válidos. Então, utilizou-se de médias móveis simples para a previsão do emprego celetista nestes municípios. Portanto, a previsão pela metodologia Box & Jenkins foi aplicada apenas nos outros doze municípios definidos no estágio de seleção de dados.

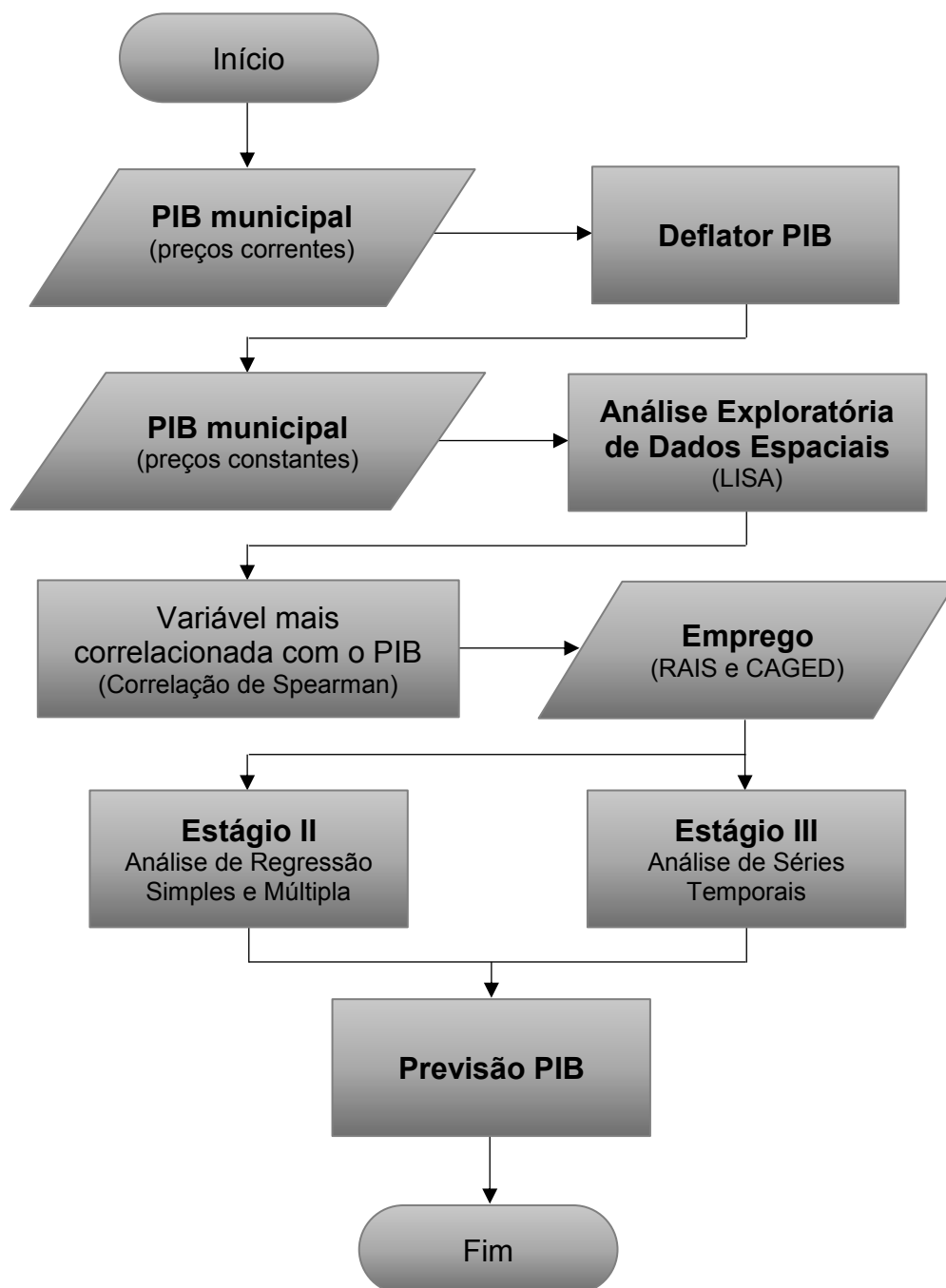
⁴ Código da função no APÊNDICE P.

⁵ Critério de informação AIC, ou *Akaike Information Criterion* é definido por $\ln \hat{\sigma}_a^2 + C \times 2/n$ em que $\hat{\sigma}_a^2$ é a estimativa da variância residual do modelo SARIMA; C é a quantidade de parâmetros estimados; e n refere-se às observações da série. O critério AIC penaliza modelos com muitas variáveis sendo que os modelos com valores menores de AIC são preferíveis (MINERVI, 2013).

As previsões foram feitas utilizando o pacote *forecast*⁶ do *software R development core team*, a partir dos modelos ajustados pela função computacional desenvolvida para este trabalho.

A FIGURA 2, a seguir, apresenta o fluxograma da metodologia proposta a fim de resumir e esclarecer as etapas mencionadas nos estágios I a III.

FIGURA 2 – FLUXOGRAMA DA METODOLOGIA PROPOSTA



FONTE: O autor (2018).

⁶ (HYNDMAN, 2014)

4 APRESENTAÇÃO DOS RESULTADOS

Os resultados da aplicação da metodologia proposta são apresentados da seguinte forma: primeiramente os resultados da Análise de Regressão para todos os municípios selecionados no estágio de tratamento e análise de dados, em seguida é apresentado o modelo SARIMA, da Análise de Séries Temporais, obtido para o município de Curitiba, de forma detalhada como exemplo, terminando, por fim, com o resultado da previsão do PIB de Curitiba e demais municípios selecionados para o ano de 2015.

4.1 RESULTADOS DA ANÁLISE DE REGRESSÃO

Os ajustes dos modelos de regressão foram feitos com os dados anuais do PIB a preços constantes (2014 = 100) e do emprego celetista obtido através da RAIS, para o período de 2002 a 2014, referentes aos 15 municípios selecionados na etapa de AEDE. Assim, estão apresentados na TABELA 3, os coeficientes estimados para 2 municípios paranaenses que foram ajustados pelo modelo descrito na equação (3.2.2.2), além da estatística t, o p-valor, e o valor do R^2 ajustado de cada modelo.

TABELA 3 – COEFICIENTES ESTIMADOS, ESTATÍSTICA T, P-VALOR E R^2 AJUSTADO, DE MODELOS DE REGRESSÃO ESTIMADOS PARA 2 MUNICÍPIOS SELECIONADOS DO ESTADO DO PARANÁ

MUNICÍPIO	COEFICIENTES ESTIMADOS	ESTATÍSTICA t	P-VALOR	R^2 AJUSTADO
Ibiporã	$\hat{\beta}_0 = 13,0127051$	(178,050)	(0,0000)	0,9564
	$\hat{\beta}_1 = 0,00016557$	(16,260)	(0,0000)	
Pinhais	$\hat{\beta}_0 = 1221664,74$	(4,031)	(0,0020)	0,9076
	$\hat{\beta}_1 = 99,627$	(10,906)	(0,0000)	

FONTE: O autor (2018).

Para o restante dos municípios foram ajustados modelos de regressão conforme apresentado na equação (3.2.2.1). Assim, a TABELA 4, adiante, apresenta os coeficientes estimados com suas respectivas estatística t e p-valor, bem como o R^2 ajustado para cada município.

TABELA 4 – COEFICIENTES ESTIMADOS, ESTATÍSTICA T, P-VALOR E R² AJUSTADO, DE MODELOS DE REGRESSÃO ESTIMADOS PARA 13 MUNICÍPIOS SELECIONADOS DO ESTADO DO PARANÁ

MUNICÍPIO	COEFICIENTES ESTIMADOS	ESTATÍSTICA t	P-VALOR	R ² AJUSTADO
Almirante Tamandaré	$\hat{\beta}_1 = 67,580$	(6,318)	(0,0001)	0,9856
	$\hat{\beta}_2 = 441826,590$	(4,627)	(0,0007)	
Arapongas	$\hat{\beta}_0 = 573064,326$	(2,933)	(0,0136)	0,9491
	$\hat{\beta}_1 = 104,968$	(14,992)	(0,0000)	
Araucária	$\hat{\beta}_1 = 195,455$	(20,490)	(0,0000)	0,9929
	$\hat{\beta}_2 = 4539395,262$	(11,880)	(0,0000)	
Cambé	$\hat{\beta}_1 = 147,753$	(70,156)	(0,0000)	0,9981
	$\hat{\beta}_2 = 267082,156$	(3,115)	(0,0098)	
Campo Largo	$\hat{\beta}_1 = 107,000$	(22,765)	(0,0000)	0,9954
	$\hat{\beta}_2 = 653691,900$	(5,887)	(0,0001)	
Carambeí	$\hat{\beta}_1 = 98,068$	(13,279)	(0,0000)	0,9949
	$\hat{\beta}_2 = 181334,400$	(2,698)	(0,0207)	
Colombo	$\hat{\beta}_0 = 1005180,308$	(4,786)	(0,0007)	0,9397
	$\hat{\beta}_1 = 75,285$	(9,367)	(0,0000)	
	$\hat{\beta}_2 = 558542,522$	(3,784)	(0,0036)	
Curitiba	$\hat{\beta}_1 = 110,225$	(37,823)	(0,0000)	0,9990
	$\hat{\beta}_2 = 8664441,664$	(4,814)	(0,0005)	
Fazenda Rio Grande	$\hat{\beta}_1 = 125,695$	(30,051)	(0,0000)	0,9937
	$\hat{\beta}_2 = 164460,833$	(3,806)	(0,0029)	
Piraquara	$\hat{\beta}_0 = -437994,10$	(-2,594)	(0,0249)	0,8239
	$\hat{\beta}_1 = 239,600$	(7,559)	(0,0000)	
Rolândia	$\hat{\beta}_1 = 87,343$	(26,114)	(0,0000)	0,9917
	$\hat{\beta}_2 = 305976,183$	(3,689)	(0,0036)	
São José dos Pinhais	$\hat{\beta}_0 = 4960214,360$	(4,159)	(0,0016)	0,9391
	$\hat{\beta}_1 = 227,250$	(13,645)	(0,0000)	
Sarandi	$\hat{\beta}_1 = 118,688$	(37,023)	(0,0000)	0,9981
	$\hat{\beta}_2 = -114800,638$	(-3,522)	(0,0048)	

FONTE: O autor (2018).

Porém, como não basta apenas um R² ajustado alto, para definir se o modelo é adequado, assim, recorreu-se à análise de resíduos de cada modelo para confirmar sua adequabilidade.

Na TABELA 5, a seguir, são mostrados os resultados dos testes de Shapiro-Wilk, teste t (para verificação de média nula), Breush-Pagan e Durbin-Watson, utilizados para a análise de resíduos para a validação dos modelos.

TABELA 5 – RESULTADOS DOS TESTES DAS ANÁLISES DOS RESÍDUOS DE MODELOS DE REGRESSÃO ESTIMADOS PARA 15 MUNICÍPIOS SELECIONADOS DO ESTADO DO PARANÁ

(continua)

MUNICÍPIO	TESTE	ESTATÍSTICA DO TESTE	P-VALOR
Almirante Tamandaré	Shapiro-Wilk	0,9558	0,6883
	Teste t	0,3564	0,7278
	Breusch-Pagan	0,0967	0,7558
	Durbin-Watson	1,5499	0,1597
Arapongas	Shapiro-Wilk	0,9795	0,9768
	Teste t	0,0000	1,0000
	Breusch-Pagan	3,4934	0,0616
	Durbin-Watson	1,6708	0,1651
Araucária	Shapiro-Wilk	0,9173	0,2308
	Teste t	0,3148	0,7583
	Breusch-Pagan	1,5524	0,2128
	Durbin-Watson	1,5998	0,1395
Cambé	Shapiro-Wilk	0,9644	0,8196
	Teste t	0,0472	0,9632
	Breusch-Pagan	0,3458	0,5565
	Durbin-Watson	1,7640	0,2453
Campo Largo	Shapiro-Wilk	0,9353	0,3994
	Teste t	0,0543	0,9576
	Breusch-Pagan	0,2537	0,6145
	Durbin-Watson	1,7082	0,2187
Carambeí	Shapiro-Wilk	0,8888	0,0940
	Teste t	0,2169	0,8319
	Breusch-Pagan	0,4901	0,4839
	Durbin-Watson	1,2783	0,0650
Colombo	Shapiro-Wilk	0,9870	0,9981
	Teste t	0,0000	1,0000
	Breusch-Pagan	3,6844	0,1585
	Durbin-Watson	1,8728	0,1867
Curitiba	Shapiro-Wilk	0,9494	0,5896
	Teste t	0,0042	0,9967
	Breusch-Pagan	0,0991	0,7529
	Durbin-Watson	1,4979	0,1154
Fazenda Rio Grande	Shapiro-Wilk	0,9318	0,3593
	Teste t	0,5391	0,5997
	Breusch-Pagan	0,7258	0,3942
	Durbin-Watson	1,7799	0,2452
Ibiporã	Shapiro-Wilk	0,9607	0,7640
	Teste t	0,0000	1,0000
	Breusch-Pagan	0,7868	0,3751
	Durbin-Watson	1,9895	0,3603
Pinhais	Shapiro-Wilk	0,9330	0,3726
	Teste t	0,0000	1,0000
	Breusch-Pagan	0,6316	0,4268
	Durbin-Watson	2,0338	0,3956

TABELA 5 – RESULTADOS DOS TESTES DAS ANÁLISES DOS RESÍDUOS DE MODELOS DE REGRESSÃO ESTIMADOS PARA 15 MUNICÍPIOS SELECIONADOS DO ESTADO DO PARANÁ

(conclusão)

MUNICÍPIO	TESTE	ESTATÍSTICA DO TESTE	P-VALOR
Piraquara	Shapiro-Wilk	0,9130	0,2018
	Teste t	0,0000	1,0000
	Breusch-Pagan	1,4508	0,2284
	Durbin-Watson	1,5710	0,1335
Rolândia	Shapiro-Wilk	0,9137	0,2063
	Teste t	-0,0830	0,9353
	Breusch-Pagan	0,0761	0,7826
	Durbin-Watson	1,3803	0,0860
São José dos Pinhais	Shapiro-Wilk	0,9580	0,7231
	Teste t	0,0000	1,0000
	Breusch-Pagan	3,4581	0,0629
	Durbin-Watson	1,3693	0,0571
Sarandi	Shapiro-Wilk	0,9735	0,9320
	Teste t	0,2368	0,8168
	Breusch-Pagan	3,3562	0,0670
	Durbin-Watson	2,0925	0,4706

FONTE: O autor (2018).

A normalidade dos resíduos é essencial para a determinação de intervalos de confiança e testes de significância (GUIMARÃES; LIRA, 2017b). Nota-se, na TABELA 5, que todos os p-valores para o teste de Shapiro-Wilk, que tem como hipótese nula a normalidade, estão acima de 0,05. Assim sendo, todos os modelos estimados têm seus resíduos normalmente distribuídos, graças a isso, os valores da estatística t e do R^2 ajustado, apresentados nas TABELAS 3 e 4, são válidos.

Entretanto, deve-se verificar também as suposições de homocedasticidade e independência dos resíduos. Para isso, é apresentado na TABELA 5 os resultados do teste de Breusch-Pagan, cuja a hipótese nula é de que todas as variâncias são estatisticamente iguais, para testar a homocedasticidade e o teste de Durbin-Watson, em que a hipótese nula é de que a autocorrelação dos resíduos é zero, significando independência dos resíduos. Nota-se que para estes testes também todos os p-valores são maiores que 0,05. Deste modo, não rejeitamos a hipótese nula nem do teste de Breusch-Pagan, nem do teste de Durbin-Watson.

Logo, por meio da TABELA 5, verifica-se que os modelos estimados são válidos, então existe a regressão entre as variáveis dependentes e independentes.

Analisando os coeficientes de determinação ajustados, apresentados nas TABELAS 3 e 4, nota-se que a medida do grau de ajustamento dos modelos de

regressão aos dados, é boa. Os modelos ajustados explicam muito bem as variações ocorridas na variável resposta, ou seja, no Produto Interno Bruto.

Através da análise estatística, observou-se que o PIB está fortemente correlacionado com o emprego (CLT) ($r_s = 0,9286$), e, também, foi possível encontrar uma relação funcional entre as duas variáveis de forma que o PIB pode ser explicado pelo comportamento do emprego celetista. Desta forma, é possível entender quanto o emprego (CLT) influencia na economia dos municípios impactando diretamente no PIB.

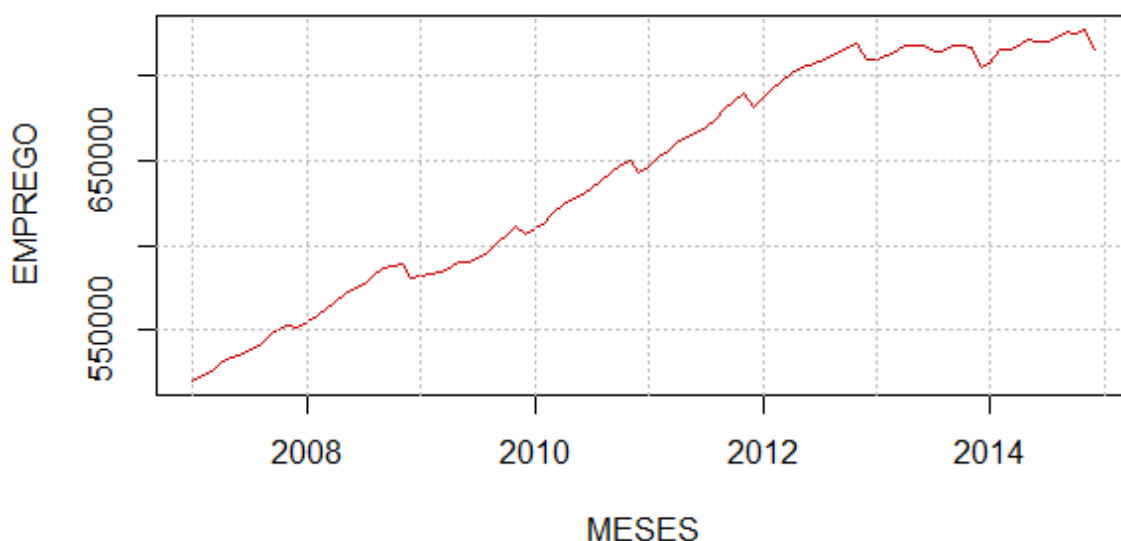
Assim, é possível concluir que os modelos ajustados poderão ser utilizados para obter a projeção do PIB. Sinaliza, também, a importância do emprego celetista para a economia de cada município integrante do presente estudo.

4.2 RESULTADOS DA ANÁLISE DE SÉRIES TEMPORAIS

Com o objetivo de ter uma apresentação dos resultados mais clara, para a Análise de Séries Temporais, resolveu-se expor nesta seção a análise detalhada feita para o município de Curitiba e Araucária, deixando os detalhes dos resultados dos demais municípios selecionados nos APÊNDICES 3 a 15.

Apresenta-se inicialmente o GRÁFICO 2 como representação gráfica da série temporal do emprego celetista do município de Curitiba, no período de janeiro de 2007 a dezembro de 2014.

GRÁFICO 2 – SÉRIE TEMPORAL DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CURITIBA - JAN. 2007- DEZ. 2014



FONTE: O autor (2018).

Apenas analisando o gráfico, consegue-se notar que existe uma sazonalidade no emprego, havendo uma pequena queda no mês de dezembro. Entretanto, como o objetivo deste trabalho é utilizar métodos estatísticos para se eximir da subjetividade, a série temporal do emprego celetista foi inserida na função *best.arima()* onde a função analisa se os dados possuem sazonalidade através do teste de Kruskal-Wallis⁷ e, se o teste indicar a presença de sazonalidade, então a *best.arima()* propõe um modelo $SARIMA(p,d,q)X(P,D,Q)_{12}$, caso contrário, ela propõe um modelo $ARIMA(p,d,q)$.

Para o município de Curitiba, o modelo ajustado pela função *best.arima()* foi um modelo $SARIMA(0,2,1)x(2,2,0)_{12}$ com AIC igual a 14,7532, com os coeficientes estimados mostrados na TABELA 6.

TABELA 6 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO $SARIMA(0,2,1)X(2,2,0)_{12}$, AJUSTADO PARA O MUNICÍPIO DE CURITIBA

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	0,0000	0,1079	0,0000
AR(2)	0,0000	0,1093	0,0000
MA(1)	- 0,5471	0,0921	0,0000
MA(2)	0,0000	0,0000	0,0000
SAR(1)	- 1,2049	0,1090	0,0000
SAR(2)	- 0,6149	0,1051	0,0000
SMA(1)	0,0000	0,0000	0,0000
SMA(2)	0,0000	0,0000	0,0000

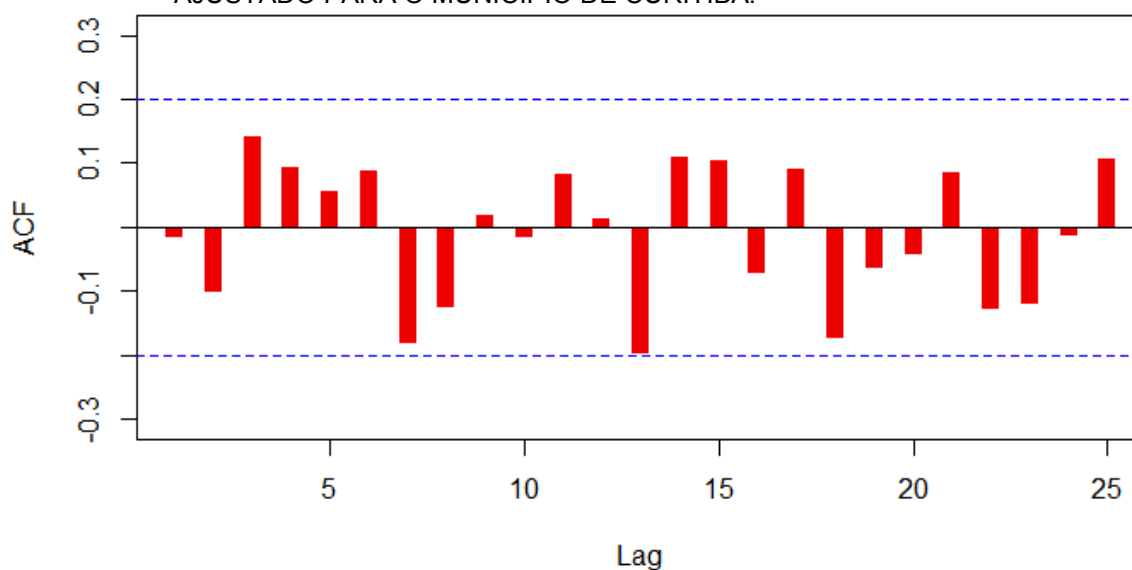
FONTE: O autor (2018).

Como dito anteriormente, ao inserir os dados da série temporal na função *best.arima()*, a função retorna o melhor modelo segundo o critério AIC com todos os testes aplicados aos resíduos válidos. Desta forma, apresentam-se os resultados dos testes da autocorrelação dos resíduos, teste de Ljung-Box, normalidade dos resíduos e periodograma acumulado.

Observando o GRÁFICO 3, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

⁷ Mais informações sobre o teste, consultar Pontes (2000).

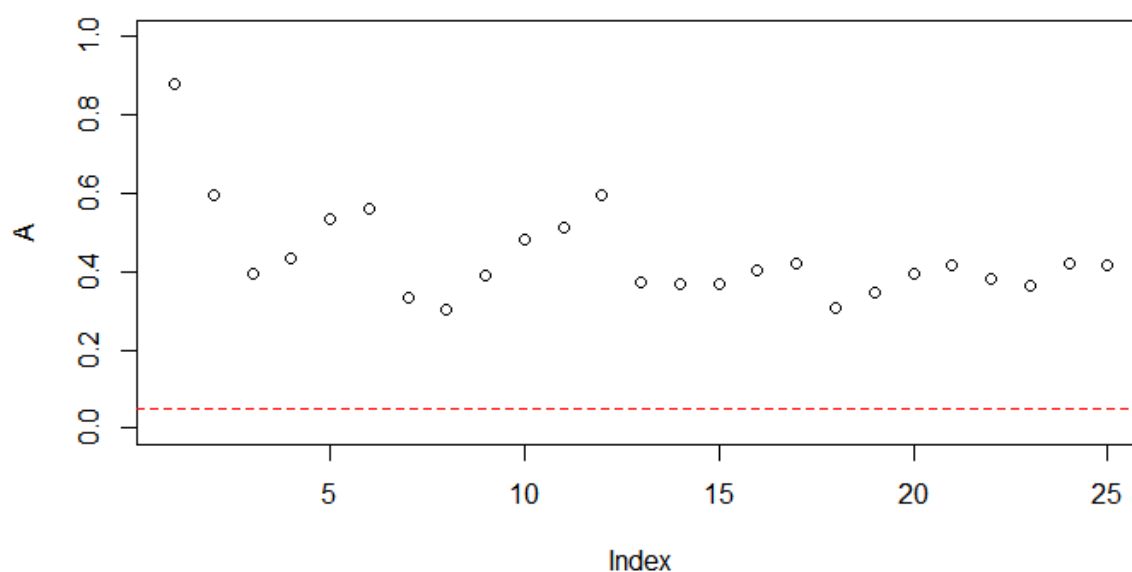
GRÁFICO 3 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO SARIMA(0,2,1)X(2,2,0)₁₂, AJUSTADO PARA O MUNICÍPIO DE CURITIBA.



FONTE: O autor (2018).

O teste de Ljung-Box apresentou o p-valor de 0,8765, não rejeitando a hipótese nula, assumindo que os resíduos são independentes e identicamente distribuídos (i.i.d). No GRÁFICO 4 abaixo, tem-se o teste de Ljung-Box aplicado individualmente para as 25 primeiras defasagens, desta forma, nota-se visualmente que em todos os lags a hipótese nula do teste foi aceita.

GRÁFICO 4 – P-VALORES DAS 25 PRIMEIRAS AUTO CORRELAÇÕES DOS RESÍDUOS DO MODELO SARIMA(0,2,1)X(2,2,0)₁₂, AJUSTADO PARA O MUNICÍPIO DE CURITIBA.



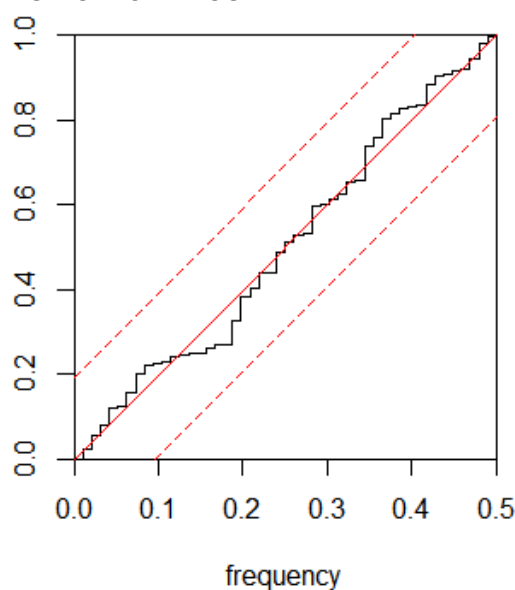
FONTE: O autor (2018).

A normalidade dos resíduos foi analisada pelo teste de Jarque-Bera. Este teste foi escolhido por ser frequentemente utilizado na econometria. A hipótese nula deste teste é a normalidade dos resíduos. O p-valor para este teste foi de 0,0585. Portanto, não foi rejeitada a hipótese de normalidade a um nível de significância de 0,05.

Por fim, o GRÁFICO 5 apresenta o periodograma acumulado dos resíduos para o modelo ajustado. Como o gráfico da série dos resíduos estimados (\hat{a}_t) se encontra dentro das linhas limites, significa que \hat{a}_t é ruído branco.

Logo, todos os testes propostos por Box, Jenkins e Reinsel (2008) foram atendidos. Logo, o modelo ajustado pela função *best.arima()* é considerado um bom modelo. Sendo assim, pode-se passar para a fase de previsão dos dados de emprego celetista para o município de Curitiba utilizando o modelo ajustado SARIMA(0,2,1)X(2,2,0)₁₂.

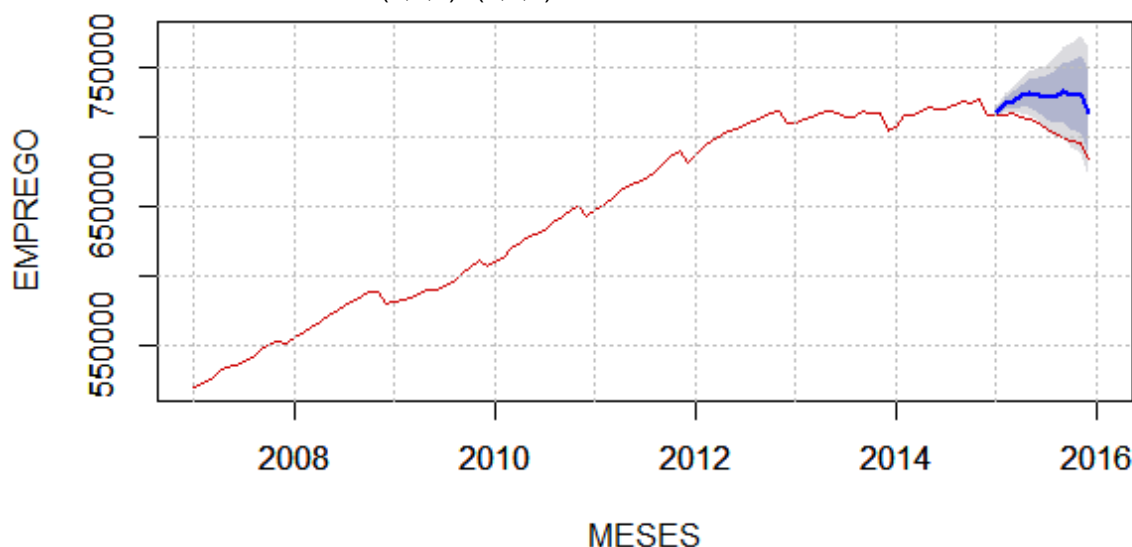
GRÁFICO 5 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO SARIMA(0,2,1)X(2,2,0)₁₂, AJUSTADO PARA O MUNICÍPIO DE CURITIBA



FONTE: O autor (2018).

Abaixo apresenta-se o GRÁFICO 6 dos valores previstos para o emprego celetista no ano de 2015 em comparação com os valores reais obtidos pela RAIS. Nota-se que ao findar do ano, os dados encontram-se dentro do intervalo de confiança de 95%, conforme a parte preenchida em cinza claro.

GRÁFICO 6 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CURITIBA PELO MODELO ARIMA(0,2,1)X(2,2,0)₁₂



FONTE: O autor (2018).

O valor real do emprego celetista em 2015 obtido pela RAIS foi de 684.142 trabalhadores regidos pela CLT. O valor previsto pelo modelo ajustado foi de 717.772. Desta forma, a diferença absoluta percentual⁸ entre o valor real e o valor projetado para o ano de 2015 é de apenas 4,69%. Entretanto, o intervalo de confiança de 95% é [670.923, 764.621], portanto, o valor real do emprego celetista obtido pelos dados da RAIS está dentro do intervalo de confiança previsto para o ano de 2015.

A TABELA 7 apresenta os modelos ajustados para cada um dos 15 municípios com os respectivos valores do critério AIC.

Nota-se que apenas para 3 municípios não foi possível encontrar um modelo da metodologia Box & Jenkins pela função *best.arima()*, utilizando para estes municípios, a metodologia de médias móveis simples como alternativa.

Sendo assim, toma-se o exemplo do município de Araucária para apresentar a previsão do emprego celetista utilizando Médias Móveis Simples.

⁸ Diferença = $\text{abs}((\text{índice observado} - \text{índice projetado}) / \text{índice observado}) * 100$

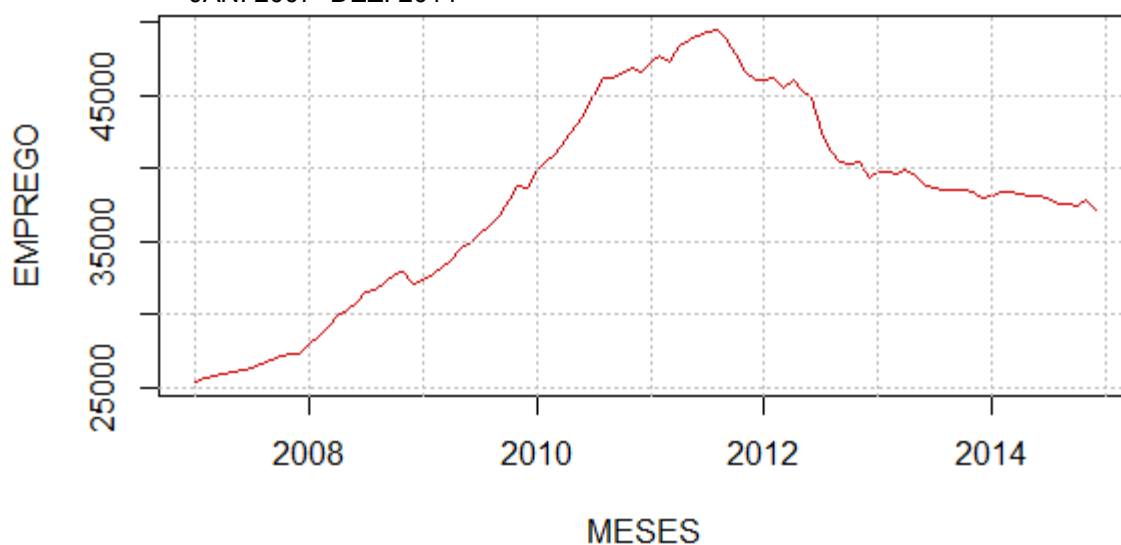
TABELA 7 – MODELOS DE PREVISÃO AJUSTADO PARA 15 MUNICÍPIOS PARANAENSES SELECIONADOS E O CRITÉRIO AIC.

MUNICÍPIO	MODELO AJUSTADO	AIC
Almirante Tamandaré	SARIMA(1,0,2)x(0,0,1) ₁₂	8,5630
Arapongas	SARIMA(0,2,1)x(0,1,1) ₁₂	9,9648
Araucária	MMS(1)	-----
Cambé	SARIMA(1,1,0)x(0,1,2) ₁₂	9,3650
Campo Largo	SARIMA(0,2,1)x(1,0,1) ₁₂	9,9400
Carambeí	SARIMA(2,0,0)x(1,2,1) ₁₂	8,2513
Colombo	SARIMA(1,2,1)x(1,0,1) ₁₂	10,7836
Curitiba	SARIMA(0,2,1)x(2,2,0) ₁₂	14,7376
Fazenda Rio Grande	MMS(1)	-----
Ibiporã	ARIMA(3,0,3)	8,5721
Pinhais	MMS(1)	-----
Piraquara	SARIMA(4,0,3)x(1,0,0) ₁₂	8,5488
Rolândia	SARIMA(4,2,0)x(2,2,1) ₁₂	10,9200
São José dos Pinhais	SARIMA(2,2,1)x(2,0,0) ₁₂	12,0652
Sarandi	SARIMA(0,2,2)x(0,1,1) ₁₂	8,2549

FONTE: O autor (2018).

O GRÁFICO 7, abaixo, apresenta a série temporal de dados do emprego celetista para a cidade de Araucária no período de análise definido na metodologia, entre janeiro de 2007 e dezembro de 2014.

GRÁFICO 7 – SÉRIE TEMPORAL DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ARAUCÁRIA - JAN. 2007- DEZ. 2014

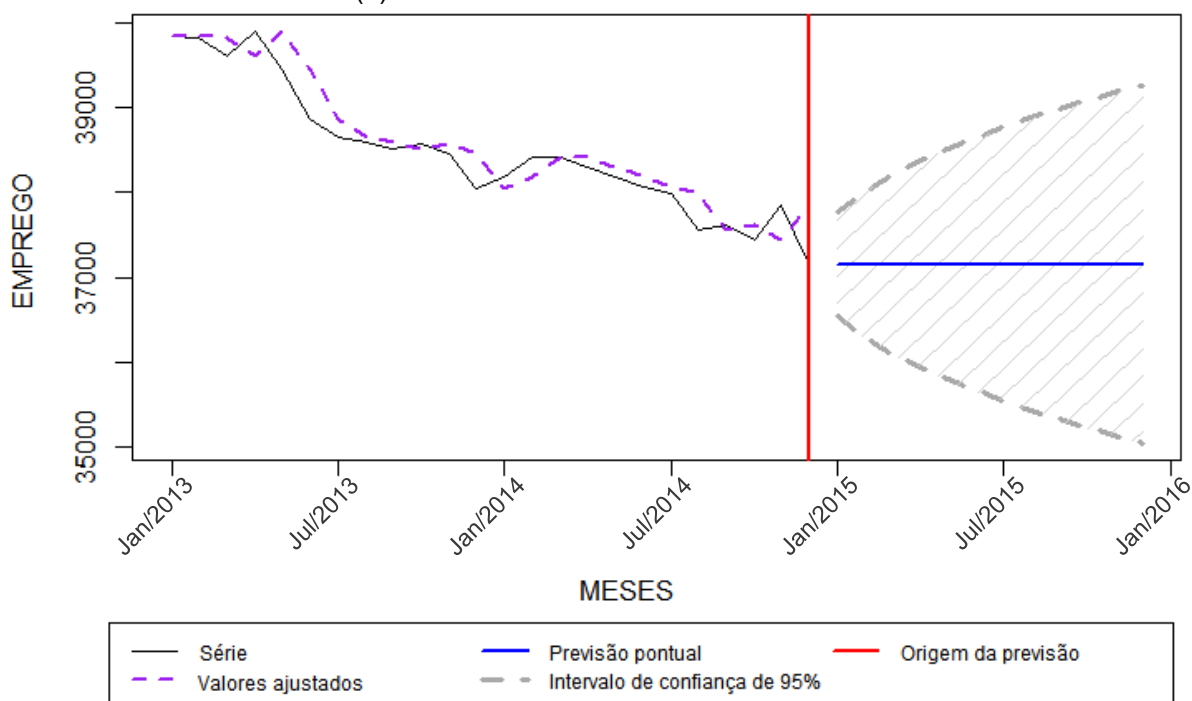


FONTE: O autor (2018).

Observa-se no GRÁFICO 7 uma queda no emprego celetista para o município de Araucária do final do ano de 2011 até o término do período selecionado. Provavelmente devido a este motivo não foi possível encontrar um modelo apropriado, pois, mesmo diferenciando para deixar a série estacionária os modelos propostos falhavam em pelo menos um dos testes aplicados aos resíduos. Desta forma, passou-se para a alternativa da MMS, cujas suposições são menos rígidas.

Como o modelo de MMS é o que apresenta o menor erro quadrático médio, feitos os testes, chegou-se a MMS(1), com erro quadrático médio de 280,6384, utilizando o período de janeiro de 2013 a dezembro de 2014. Ou seja, a previsão será o último valor da série. O teste de Jarque-Bera aplicado aos resíduos apresentou p-valor de 0,8266 e o teste de Ljung-Box, também aplicado aos resíduos, apresentou o p-valor de 0,4777. Sendo assim, o GRÁFICO 8 apresenta a previsão com horizonte de tamanho 12 a fim de prever o emprego celetista para o ano de 2015.

GRÁFICO 8 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ARAUCÁRIA PELO MODELO MMS(1)



FONTE: O autor (2018).

O valor real do emprego celetista para o município de Araucária no ano de 2015 obtido pela RAIS foi de 37.370. O valor previsto pelo modelo ajustado foi de 37.155. Desta forma, a diferença absoluta percentual entre o valor real e o valor

projetado para o ano de 2015 é de apenas 0,58%, sendo o intervalo de confiança de [35049, 39261], a um nível de 95%.

4.3 PREVISÃO DO PIB DE CURITIBA PARA O ANO DE 2015

Como a metodologia proposta é de prever o PIB municipal através do emprego celetista, utilizando o modelo de regressão ajustado no estágio II da pesquisa, aplica-se o valor previsto, juntamente com os valores do intervalo de confiança, para prever o valor do PIB de Curitiba no ano de 2015.

Sendo assim, o valor previsto para o PIB de Curitiba está entre R\$ 73.935.714.600,00 e R\$ 84.261.234.200,00, sendo a previsão pontual de R\$ 78.954.975,00. O PIB a preço corrente no ano de 2015 foi de R\$ 83.864.936.000,00, entretanto, deve-se lembrar que a metodologia foi utilizada com o PIB a preços constantes (2014=100), desta forma, utilizando o deflator, tem-se que o PIB de Curitiba a preços constantes (2014=100) no ano de 2015 foi de R\$ 75.779.286.610,00. O valor previsto encontra-se dentro do intervalo de confiança de 95%, obtendo uma diferença absoluta percentual entre o valor real e o valor projetado para o ano de 2015 de apenas 4,19%.

O motivo de ter sido feita a previsão do PIB apenas para o ano de 2015 é que os dados do PIB municipal estão disponíveis até este período, para uso de comparação com os dados reais. O horizonte de projeção até pode ser aumentado no *software R Development Core Team*, mas o intervalo de confiança para a projeção se torna muito grande, não sendo seguro para uma projeção para períodos muito longo.

4.4 PREVISÃO DO PIB PARA OS MUNICÍPIOS SELECIONADOS

A TABELA 8 abaixo, apresenta a previsão pontual para o ano de 2015 para os municípios selecionados, bem como os limites do intervalo de previsão, o valor real obtido através do IBGE e, ainda, a diferença absoluta percentual entre o valor real e o valor projetado.

TABELA 8 – PREVISÃO DO PIB MUNICIPAL UTILIZANDO A METODOLOGIA PROPOSTA EM COMPARAÇÃO COM O VALOR DO PIB A PREÇOS CONSTANTES (2014=100) (R\$ 1.000), UTILIZANDO INTERVALO DE CONFIANÇA DE 99% (OU 95% ???)

MUNICÍPIO	LIMITE INFERIOR DE PREVISÃO	PREVISÃO PONTUAL	LIMITE SUPERIOR DE PREVISÃO	PIB (IBGE)	DIFERENÇA (%)
Almirante Tamandaré	1.092.183,40	1.165.283,33	1.238.382,59	1.132.661,97	2,88
Arapongas	3.921.207,63	4.232.046,22	4.542.884,81	3.513.376,71	20,46
Araucária	11.389.938,60	11.801.525,79	12.213.112,97	12.607.074,18	6,39
Cambé	2.974.828,75	3.273.239,58	3.571.651,88	3.199.035,87	2,32
Campo Largo	2.708.010,57	2.949.713,94	3.191.416,24	3.243.722,78	9,06
Carambeí	1.064.545,20	1.147.055,89	1.229.566,48	1.071.246,05	7,08
Colombo	3.999.352,18	4.213.531,23	4.427.710,28	3.968.657,27	6,17
Curitiba	72.329.920,56	78.954.975,00	85.903.038,09	75.779.286,61	4,19
Fazenda Rio Grande	1.528.076,87	1.666.516,08	1.804.955,30	1.589.442,49	4,85
Ibiporã	1.829.324,49	2.140.483,50	2.504.569,11	2.349.704,53	8,90
Pinhais	5.225.192,67	5.484.305,56	5.743.418,45	4.599.835,55	19,23
Piraquara	750.370,87	909.855,33	1.069.339,80	954.942,62	4,72
Rolândia	1.488.094,69	2.079.458,33	2.670.822,85	2.073.396,58	0,29
São José dos Pinhais	21.381.653,87	23.684.307,67	25.986.959,20	20.581.378,87	15,08
Sarandi	1.060.884,97	1.247.872,58	1.434.859,59	1.198.836,18	4,09

FONTE: O autor (2018).

Comparando a TABELA 8 com as TABELAS 3 e 4, observa-se que os municípios que obtiveram a maior diferença absoluta percentual entre o valor real e o valor projetado, foram os municípios que obtiveram menor R^2 ajustado, com exceção de Piraquara que, mesmo tendo o valor do R^2 ajustado menor que todos os outros, o erro de previsão foi o quarto menor.

Muitas são as possibilidades de os erros de previsão serem tão grandes, como por exemplo, por não terem sido considerados os erros de regressão, as séries de dados analisadas são curtas, o horizonte de projeção da série temporal é longo e, ainda, há o caso da aleatoriedade.

5 CONCLUSÃO

A presente pesquisa teve como objetivo propor uma nova metodologia para a previsão do PIB municipal anual através de outra variável, com uma série temporal maior, que fosse fortemente correlacionada com o mesmo. Sendo assim, procurou-se definir qual seria a variável mais correlacionada com o PIB municipal a preços constantes; identificar *clusters* espaciais do PIB municipal no estado do Paraná em 2013; encontrar uma relação funcional entre o PIB municipal e a variável mais correlacionada com ele; e ajustar modelos de previsão para a variável mais correlacionada com o PIB municipal a preços constantes para cada município com correlação positiva nas aglomerações espaciais encontradas.

A metodologia da pesquisa foi dividida em três estágios, sendo o primeiro estágio definido como tratamento e seleção de dados, utilizando um deflator cedido pelo IPARDES a fim de transformar os valores do PIB nominal municipal em valores do PIB a preços constantes, selecionar a variável mais correlacionada com o PIB e selecionar os municípios a serem empregues a análise de regressão e de séries temporais. No segundo estágio foi realizada uma análise de regressão simples ou múltipla para cada município selecionado no estágio anterior. O terceiro estágio se caracterizou por ser realizada uma análise nas séries temporais do emprego celetista municipal dos 15 municípios selecionados no primeiro estágio.

No primeiro estágio foi identificada que a variável com maior correlação com o PIB municipal a preços constantes (2014 = 100) para o estado do Paraná foi o emprego celetista, obtido através da RAIS, apresentando um coeficiente de correlação de Spearman com valor 0,9286, indicando uma correlação muito forte entre as duas variáveis. Além disso, foi realizada uma AEDE utilizando o LISA para identificar os *clusters* espaciais do PIB real municipal no ano de 2013, sendo esta data escolhida por apresentar o maior valor do PIB a preços constantes (2014 = 100) do estado do Paraná. A utilização desta técnica foi com objetivo de selecionar alguns municípios do Estado do Paraná para compor o presente estudo. Foram identificados 15 municípios paranaenses com correlação H-H, sendo eles: Almirante Tamandaré, Arapongas, Araucária, Cambé, Campo Largo, Carambeí, Colombo, Curitiba, Fazenda Rio Grande, Ibiporã, Pinhais, Piraquara, Rolândia, São José dos Pinhais e Sarandi.

Buscando encontrar uma relação funcional entre emprego celetista e PIB municipal através da regressão simples ou múltipla, observou-se, no segundo estágio,

que os modelos estimados apresentaram valores altos dos coeficientes de determinação ajustados, mostrando assim que os mesmos explicam muito bem as variações ocorridas PIB. Os modelos ainda foram avaliados a um nível de significância de 5% para validar a regressão. Desta forma, foi possível concluir que os modelos ajustados podem fornecer uma estimativa do PIB baseados no emprego celetista devido ao seu grau de ajustamento elevado.

Enfim o terceiro estágio expôs os modelos de previsão ajustados para o emprego celetista dos municípios definidos no primeiro estágio, tendo como prioridade encontrar modelos definidos pela metodologia Box & Jenkins, sendo 12 ajustados por esta maneira, e, como alternativa, ajustar modelos de MMS para os municípios que não puderam ser ajustados pela metodologia Box & Jenkins, sendo 3 no total.

Ao fim desses três estágios, foi apresentada a previsão do PIB a preços constantes para o ano de 2015 de cada município selecionado no primeiro estágio, utilizando a metodologia proposta. Em sua maioria, os valores previstos se aproximaram dos valores reais, estando dentro do intervalo de confiança de 95% para a previsão.

Deve-se salientar que a metodologia empregada não pode ser generalizada a todos os 399 municípios do Paraná, pois a seleção dos 15, analisados neste trabalho, foi com o objetivo de reduzir o número deles e não de obter uma amostra representativa de municípios.

6 CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

A proposta de previsão do PIB através do emprego celetista se mostrou coesa com os resultados reais dentro do intervalo de confiança para o ano de 2015. Devido ao horizonte de projeção ser de tamanho 12, o intervalo de confiança para o modelo ajustado ficou amplo, porém todos os valores pontuais foram aproximados do valor real observado. Mesmo com uma série curta de dados anuais para o PIB e emprego celetista, a metodologia mostrou-se eficiente e consistente. Devido a isso, a previsão do PIB municipal foi possível aplicando técnicas estatísticas aos dados, sem qualquer subjetividade.

Como os intervalos de previsão ficaram largos, faz-se necessário prosseguir o estudo, principalmente com séries estatísticas mais longas do PIB municipal, o que hoje não é possível, e com o decorrer dos anos, será possível se ter séries mais longas. Os modelos tanto de regressão quanto de séries temporais devem ser estimados a cada nova informação, oferecendo assim maior segurança na sua utilização.

A função *best.arima()* desenvolvida para este trabalho pode ser ainda melhorada, fazendo com que a obtenção de um modelo ARIMA, válido em todos os testes propostos por Box, Jenkins e Reinsel (2008), seja ainda mais rápido.

REFERÊNCIAS

- ANSELIN, L. Local indicators of spatial association – LISA, **Geographical analysis**, Ohio, v. 27, n. 2, p. 93-115, 1995. Disponível em: <http://dces.wisc.edu/wp-content/uploads/sites/30/2013/08/W4_Anselin1995.pdf>. Acesso em: 5 fev. 2018.
- BACHA, E. Saída para a crise tem mão dupla, **Estudos Avançados**, São Paulo, v. 31, n. 89, p. 23-27. 2017. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142017000100023&lng=pt&nrm=iso>. Acesso em: 5 fev. 2018.
- BASTOS, E. K. X.; ARAUJO, B. C. **Variações na renda real não captadas pelo PIB: efeito termos de troca e efeito preços relativos entre comerciáveis e não comerciáveis: Brasil, 1991-2016**, IPEA: Carta de Conjuntura, n. 34, 2017. Disponível em: <http://www.ipea.gov.br/portal/images/stories/PDFs/conjuntura/170418_cc34_mercado-de-trabalho.pdf>. Acesso em: 5 fev. 2018.
- BLANCHARD, O. **Macroeconomia**, 5. ed., São Paulo: Pearson Prentice Hall. 2012. 688 p.
- BOX, G.E.; JENKINS, G.M.; REINSEL, G.C. **Time series analysis: forecast and control**. 4. ed., New Jersey: John Wiley & Sons, 2008. 756 p.
- BRASIL, Ministério do Trabalho e Emprego, **Programa de Disseminação das Estatísticas do Trabalho**, Brasília, 2017. Disponível em: <<http://pdet.mte.gov.br/microdados-rai-e-caged>>. Acesso em: 5 fev. 2018.
- BUENO, R.L.S. **Econometria de séries temporais**. 2ª ed. São Paulo: Cengage Learning, 2011.
- CARVALHO, S. S. Construção de séries longas de alta frequência de indicadores do mercado de trabalho com a PME e a PNADC. **Carta de Conjuntura do Ipea**, v. 32, 2016. Disponível em: <http://repositorio.ipea.gov.br/bitstream/11058/7255/1/cc32_nt_construcao_series_longas_alta.pdf>. Acesso em: 5 fev. 2018.
- CHARNET, R.; FREIRE, C.; CHARNET, E.; BONVINO, H. **Análise de Modelos de Regressão Linear com aplicações**. Campinas: Unicamp, 1999. 365 p.
- COYLE, D. **GDP: a brief but affectionate history**. Princeton University Press, Princeton and Oxford. 2014.
- FEIJÓ, C.A.; RAMOS, R.L.O. **Contabilidade social: o novo sistema de contas nacionais do Brasil**. 4ª ed. Rio de Janeiro: Campus. 2013. 391 p.
- FERREIRA, R. N.; MATOS, R., **Dinamismo do mercado de trabalho formal e mobilidade espacial de trabalhadores**, In: XVI Encontro Nacional de Estudos Populacionais, 2008 Caxambú. **Anais...** Caxambú: ABEP, 2008. Disponível em: <<http://www.abep.org.br/publicacoes/index.php/anais/article/view/1716/1676>>. Acesso em: 5 fev. 2018.

FIGUEIREDO FILHO, D. et al. O que Fazer e o que Não Fazer com a Regressão: pressupostos e aplicações do modelo linear de Mínimos Quadrados Ordinários (MQO). **Revista Política Hoje**, v. 20, n. 1, p. 44-99, 2011. Disponível em: <<https://periodicos.ufpe.br/revistas/politica hoje/article/view/3808/3112>>. Acesso em: 5 fev. 2018.

GARCÍA, J. F.; URDANIVIA, C. A. D. Para uma breve historia de la econometría, **Política y Cultura**, México D.F.: Universidad Autónoma Metropolitana, n. 13, p. 7-32, 2000. Disponível em: <<http://www.redalyc.org/articulo.oa?id=26701302>>. Acesso em: 5 fev. 2018.

GUIMARÃES, I. M.; LIRA, S. A. **Ajuste de Modelos De Regressão: Produto Interno Bruto e Emprego Formal, Segundo Municípios Selecionados Do Estado Do Paraná**. Proceedings of the XXXVIII Iberian Latin American Congress on Computational Methods in Engineering, 2017a. Disponível em: <<http://dx.doi.org/10.20906/CPS/CILAMCE2017-0083>>. Acesso em: 5 fev. 2018.

GUIMARÃES, I. M.; LIRA, S. A. **Modelos de regressão: produto interno bruto X emprego celetista, segundo municípios selecionados do Estado do Paraná**. In: SIMPÓSIO DE MÉTODOS NUMÉRICOS EM ENGENHARIA, 2., 2017, Curitiba, **Anais...** Curitiba: UFPR. 2017b. Disponível em: <<http://eventos.ufpr.br/smne/SMNE2017/paper/view/558/220>>. Acesso em: 5 fev. 2018.

GUJARATI, D.; PORTER, D. C. **Econometria Básica**. 5 ed. Porto Alegre: AMGH, 2011.

HOBBSAWM, Eric. **Sobre história**. São Paulo: Companhia das Letras, 2013.

HYNDMAN, R. J. et all. **forecast: Forecasting functions for time series and linear models**. R package version 5.3. 2014. <https://CRAN.R-project.org/package=forecast>.

IBGE. **Sistema de Contas Nacionais**. Diretoria de pesquisas. Disponível em: <<https://www.ibge.gov.br/estatisticas-novoportal/economicas/contas-nacionais/9088-produto-interno-bruto-dos-municipios.html?&t=series-historicas>>. Acesso em: abr. 2017.

IBGE. **Sistema de Contas Nacionais – Brasil: Referência 2010**. Diretoria de pesquisas. 2015. Disponível em: <ftp://ftp.ibge.gov.br/Contas_Nacionais/Sistema_de_Contas_Nacionais/Notas_Metodologicas_2010/02_estrutura_scn.pdf>. Acesso em: 5 fev. 2018.

IPARDES. **PIB do Paraná encolhe menos do que o País e recessão perde força no Estado**, Curitiba, 2016. Disponível em: <http://www.ipardes.gov.br/index.php?pg_conteudo=1&cod_noticia=801>. Acesso em: 5 fev. 2018.

KUZNETS, S. **National Income, 1929-1932**. 73rd US Congress, 2d session, Senate document, n. 124. Vol. 7. 1934. Disponível em: <https://fraser.stlouisfed.org/scribd/?title_id=971&filepath=/files/docs/publications/national_income_1934/19340104_nationalinc.pdf>. Acesso em: 5 fev. 2018.

LEOTTI, V. B., COSTER, R., RIBOLDI, João. Normalidade de variáveis: métodos de verificação e comparação de alguns testes não-paramétricos por simulação. **Revista HCPA**. Porto Alegre. Vol. 32, no. 2 (2012), p. 227-234, 2012. Disponível em: <<http://hdl.handle.net/10183/158102>>. Acesso em: 5 fev. 2018.

LIEBEL, M. J., **Previsão de Receitas Tributárias – O caso do ICMS no Estado do Paraná**. 2004. 118 p. Dissertação (Mestrado Profissionalizante em Engenharia) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/4711/000459140.pdf>>. Acesso em: 5 fev. 2018.

LIRA, S. A., **Análise de correlação: abordagem teórica e de construção dos coeficientes com aplicações**. 2004. 196 p. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setores de Ciências Exatas e de Tecnologia, Universidade Federal do Paraná, Curitiba, 2004. Disponível em: <http://www.ipardes.gov.br/biblioteca/docs/dissertacao_sachiko.pdf>. Acesso em: 5 fev. 2018.

MICHELS, E., OLIVEIRA, N., WOLLENHAUPT, S. **Fundamentos da Economia**. Curitiba: InterSaberes. 2013.

MINERVI, N. A. **Análise fatorial e de séries temporais aplicadas às séries dos acidentes de trânsito nas rodovias federais brasileiras**. 2013. 316 p. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setores de Ciências Exatas e de Tecnologia, Universidade Federal do Paraná, Curitiba, 2013. Disponível em: <<http://acervodigital.ufpr.br/bitstream/handle/1884/31902/R%20-%20D%20-%20NIVALDO%20APARECIDO%20MINERVI.pdf?sequence=1f>>. Acesso em: 5 fev. 2018.

MONTGOMERY, D.C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. 5 ed. New Jersey: John Wiley & Sons, 2012. 645 p.

MORENO, C., CHASSÉ, D. S., FUHR, L. **A Métrica do Carbono: Abstrações Globais e Epistemicídio Ecológico**. Rio de Janeiro: Fundação Heinrich Böll. 2016. Disponível em: <https://br.boell.org/sites/default/files/carbonmetrics_livro_boll.pdf>.

MORETIN, P. A.; TOLOI, C.M.C. **Modelos para previsão de séries temporais**. 2. ed. São Paulo: Edgard Blücher, 2006. 538 p.

MOURA, R., LIRA, S. A., & CINTRA, A. P. de U. Arranjos espaciais: concentração e mobilidade que redesenham aglomerações e centros. **Caderno IPARDES**. Curitiba, v.2, n.2, p. 51-67. 2012. Disponível em: <<http://www.ipardes.pr.gov.br/ojs/index.php/cadernoipardes/article/view/536/718>>. Acesso em: 5 fev. 2018.

MOURA, R.; LIRA, S. A. Aplicação da análise exploratória espacial na identificação de configurações territoriais. **Revista Brasileira de Estudos Populacionais**. Rio de Janeiro, v. 28, n.1, p. 153-168. 2011. Disponível em: <<http://dx.doi.org/10.1590/S0102-30982011000100008>>. Acesso em: 5 fev. 2018.

PEROBELLI, F. S.; FERREIRA, P. G.; FARIA, W. R. **Análise de convergência espacial do PIB per-capita em Minas Gerais: 1975-2003**. In: XI Encontro Regional

de Economia, 2006, Fortaleza. Nordeste: Estratégias de Desenvolvimento Regional. Fortaleza: BNB, 2006. Disponível em: <http://edi.bnb.gov.br/content/aplicacao/Eventos/forumbnb2006/docs/analise_de_convergencia.pdf>. Acesso em: 5 fev. 2018.

PHILIPSEN, D. **The little big number: how GDP came to rule the world and what to do about it**. Princeton: Princeton University Press. 2015.

PONTES, A. C. F. **Obtenção dos níveis de significância para os testes de Kruskal-Wallis, Friedman e comparações múltiplas não-paramétricas**. 2000. Dissertação (Mestrado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2000. Disponível em: <<http://dx.doi.org/10.11606/D.11.2000.tde-15032002-093020>>. Acesso em: 5 fev. 2018.

R Core Team (2016). R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. URL <https://www.R-project.org/>.

RAMOS, C. A. O emprego nos anos 90: o regime macroeconômico importa? **Mercado de Trabalho – Conjuntura e Análise**, n. 22, IPEA, nov. 2003. Disponível em: <http://ipea.gov.br/agencia/images/stories/PDFs/mercadodetrabalho/mt_22f.pdf>. Acesso em: 5 fev. 2018.

RIBAS, R. P.; SOARES, S. S. D. O atrito nas pesquisas longitudinais: o caso da pesquisa mensal de emprego (PME/IBGE). **Estudos Econômicos**, São Paulo, v. 40, n. 1, p. 213-244, mar. 2010. Disponível em: <<http://dx.doi.org/10.1590/S0101-41612010000100008>>. Acesso em: 5 fev. 2018.

ROMERO, J. A. R. Análise espacial da pobreza municipal no estado de minas gerais-1991-2000. In: XIV Encontro Nacional de Estudos Populacionais, 2006 Caxambú. **Anais...** Caxambú: ABEP, 2006. Disponível em: <<http://www.abep.org.br/publicacoes/index.php/anais/article/viewFile/1531/1495>>. Acesso em: 5 fev. 2018.

SABÓIA, J. Emprego industrial no Brasil: situação atual e perspectivas para o futuro. **Revista de Economia Contemporânea**, v.5, Edição Especial, Rio de Janeiro: IE/UFRJ, 2001, p.207-229. Disponível em: <http://www.ie.ufrj.br/images/pesquisa/publicacoes/rec/REC%205/REC_5.Esp_09_Emprego_industrial_no_brasil_situacao_atual_e_perspectivas_para_o_futuro.pdf>. Acesso em: 5 fev. 2018.

SILVA, J. F. S. et al. Correlação entre produção agrícola, variáveis clínicas-demográficas e câncer de próstata: um estudo ecológico. **Ciência & Saúde Coletiva**, v. 20, p. 2805-2812, 2015. Disponível em: <<http://www.fucamp.edu.br/editora/index.php/racg/article/view/897/656>>. Acesso em: 5 fev. 2018.

SILVA, P. M. S. **Mercado de capitais: sistema protetivo dos interesses coletivos dos investidores e consumidores**. 2008. 165 p. Dissertação (Mestrado em Direito), Universidade de Ribeirão Preto, Ribeirão Preto, 2008. Disponível em:

<<http://tede.unaerp.br:8180/bitstream/tede/26/1/PLINIO%20MARCOS%20DE%20SOUSA%20SILVA.pdf>>. Acesso em: 5 fev. 2018.

STEFFEN, D. **Intervalo “Bootstrap” Para Previsões De Séries Temporais Obtidas Pelo Método Theta**. 2010. 116 p. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setores de Ciências Exatas e de Tecnologia, Universidade Federal do Paraná, Curitiba, 2010. Disponível em: <http://acervodigital.ufpr.br/bitstream/handle/1884/24294/DISSERTACAO_DANIEL_STEFFEN.pdf?sequence=1>. Acesso em: 5 fev. 2018.

APÊNDICE 1 – CORRELAÇÃO DE SPEARMAN ENTRE PIB DOS MUNICÍPIOS E ALGUMAS OUTRAS VARIÁVEIS

TABELA A.1.1 - CORRELAÇÃO DE SPEARMAN ENTRE PIB A PREÇOS
CONSTANTES (2014 = 100) DOS MUNICÍPIOS DO PARANÁ
SEGUNDO ALGUMAS VARIÁVEIS

VARIÁVEL	COEFICIENTE DE CORRELAÇÃO
Emprego Formal	0,92366
Emprego Celetista	0,92857
Consumo de energia elétrica total	0,92369
Consumo de energia elétrica comercial	0,92473
Consumo de energia elétrica industrial	0,79821
Rendimento médio mensal	0,55402

FONTE: O autor (2018).

APÊNDICE 2 – RESULTADOS DOS TESTES DE NORMALIDADE E LJUNG-BOX DOS RESÍDUOS PARA OS MODELOS DE PREVISÃO AJUSTADOS

TABELA A.2.1 - RESULTADOS DO TESTE DE JARQUE-BERA E LJUNG-BOX PARA OS 15
RESÍDUOS DOS MODELOS DE PREVISÃO AJUSTADOS NO ESTÁGIO III

MUNICÍPIO	MODELO	JARQUE-BERA (P-VALOR)	LJUNG-BOX (P-VALOR)
Almirante Tamandaré	SARIMA(1,0,2)x(0,0,1) ₁₂	0,7843	0,1995
Arapongas	SARIMA(0,2,1)x(0,1,1) ₁₂	0,7831	0,2115
Araucária	MMS(1)	0,8266	0,4777
Cambé	SARIMA(1,1,0)x(0,1,2) ₁₂	0,1650	0,2070
Campo Largo	SARIMA(0,2,1)x(1,0,1) ₁₂	0,3359	0,1896
Carambeí	SARIMA(2,0,0)x(1,2,1) ₁₂	0,8530	0,5506
Colombo	SARIMA(1,2,1)x(1,0,1) ₁₂	0,4261	0,6837
Curitiba	SARIMA(0,2,1)x(2,2,0) ₁₂	0,0585	0,8765
Fazenda Rio Grande	MMS(1)	0,4000	0,9790
Ibiporã	ARIMA(3,0,3)	0,2546	0,6488
Pinhais	MMS(1)	0,2434	0,0887
Piraquara	SARIMA(4,0,3)x(1,0,0) ₁₂	0,1269	0,6001
Rolândia	SARIMA(4,2,0)x(2,2,1) ₁₂	0,2591	0,3219
São José dos Pinhais	SARIMA(2,2,1)x(2,0,0) ₁₂	0,4142	0,9190
Sarandi	SARIMA(0,2,2)x(0,1,1) ₁₂	0,4104	0,6877

FONTE: O autor (2018).

APÊNDICE 3 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ

Apresenta-se, abaixo, a TABELA A.3.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de Almirante Tamandaré. Observa-se que todos os coeficientes são significativos a um nível de significância de 0,5%.

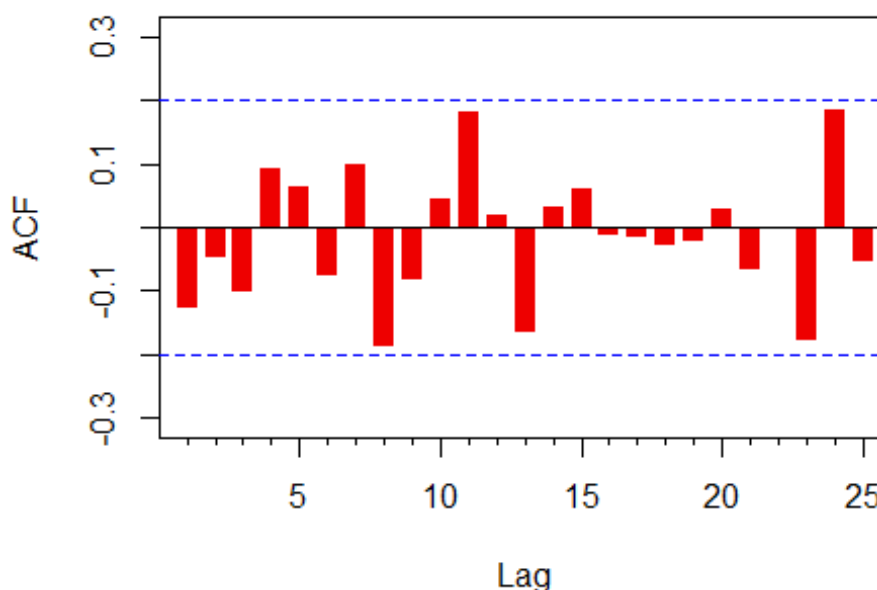
TABELA A.3.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	0,9952	0,0062	0,0000
MA(1)	0,4619	0,1116	0,0000
MA(2)	0,4293	0,0991	0,0000
SAR(1)	0,0000	0,0000	0,0000
SAR(2)	0,0000	0,0000	0,0000
SMA(1)	0,2760	0,0876	0,0016
SMA(2)	0,0000	0,0000	0,0000
INTERCEPTO	9045,134	1529,423	0,0000

FONTE: O autor (2018).

Observando o GRÁFICO A.3.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

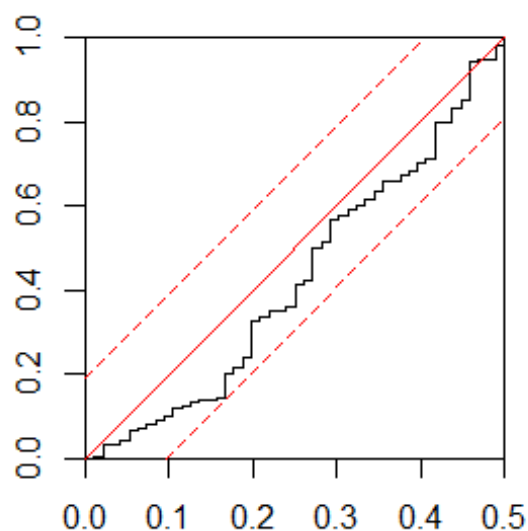
GRÁFICO A.3.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.3.2. Devido ao fato do GRÁFICO da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

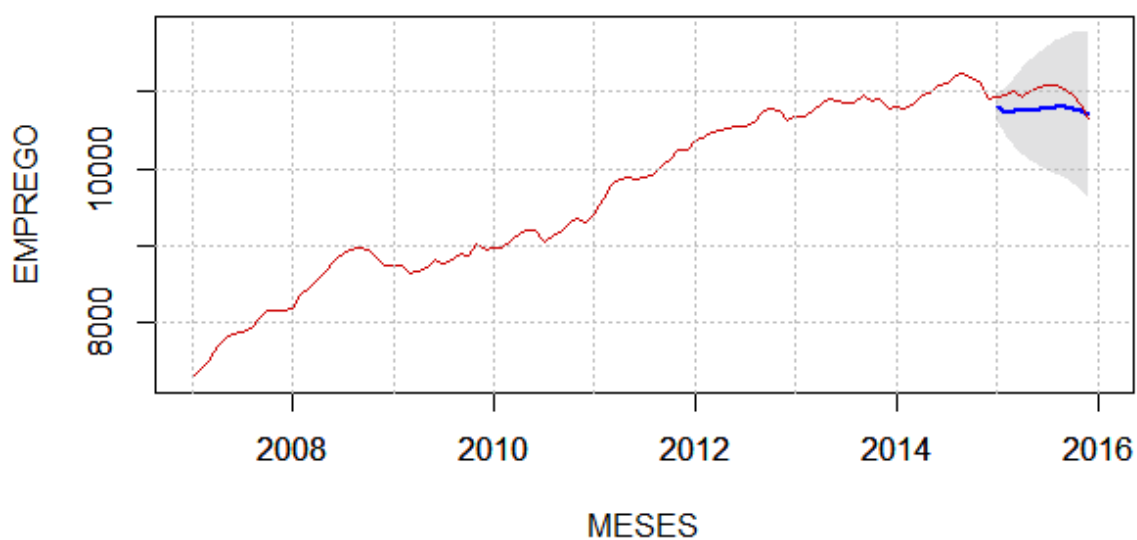
GRÁFICO A.3.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.3.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.3.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ALMIRANTE TAMANDARÉ PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 4 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE ARAPONGAS

Apresenta-se, abaixo, a TABELA A.4.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arma()* para o município de Arapongas. Observa-se que todos os coeficientes são significativos a um nível de significância de 0,5%.

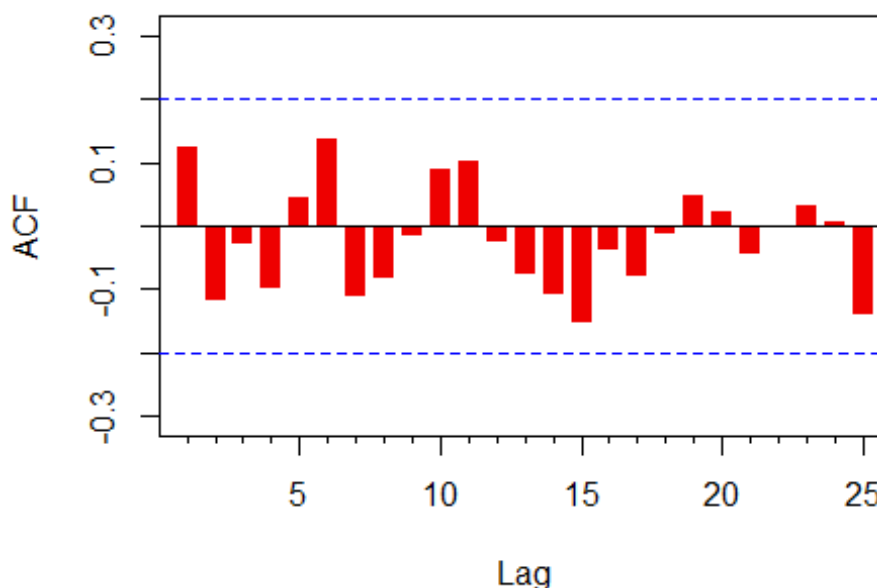
TABELA A.4.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ARAPONGAS

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)			
MA(1)			
MA(2)	-0,8125	0,0865	0,0000
SAR(1)			
SAR(2)			
SMA(1)			
SMA(2)	-0,7680	0,1634	0,0000
INTERCEPTO			

FONTE: O autor (2018).

Observando o GRÁFICO A.4.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

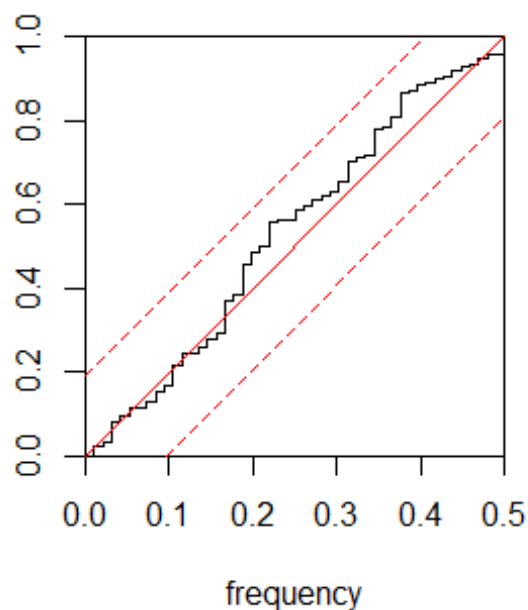
GRÁFICO A.4.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ARAPONGAS



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.4.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

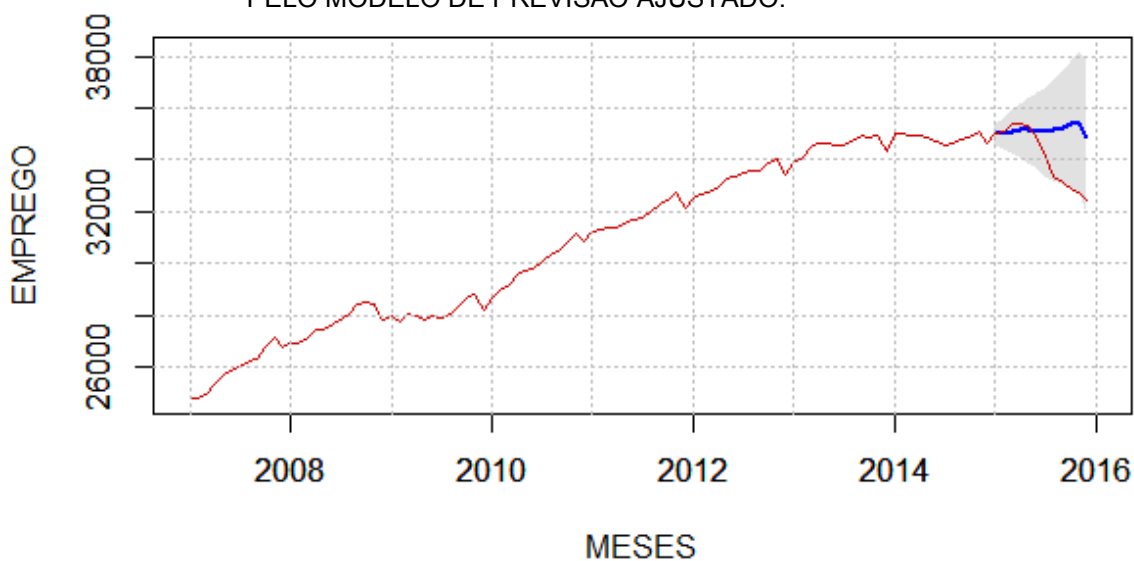
GRÁFICO A.4.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ARAPONGAS



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.4.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.4.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ARAPONGAS PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 5 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE CAMBÉ

Apresenta-se, abaixo, a TABELA A.5.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de Cambé. Observa-se que todos os coeficientes são significativos a um nível de significância de 5%.

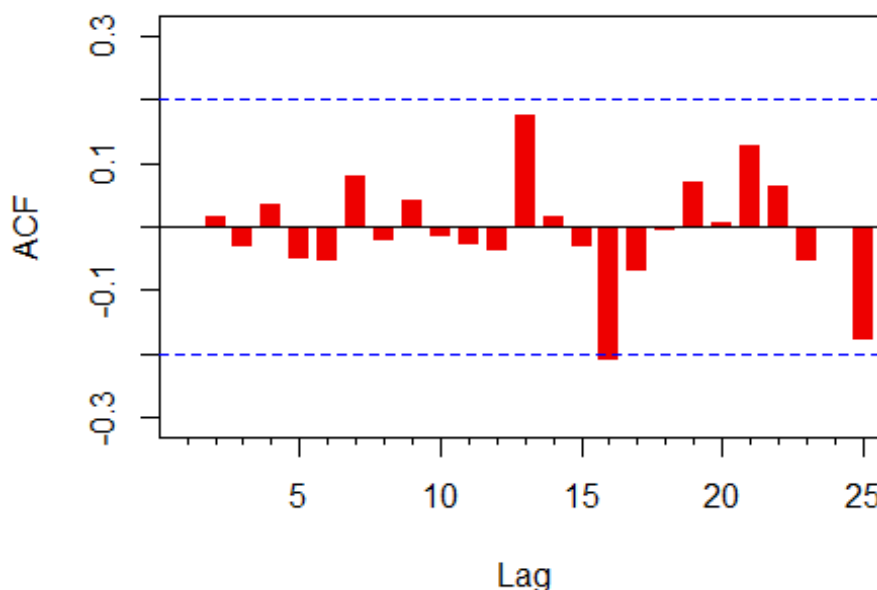
TABELA A.5.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMBÉ

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	0,5438	0,1058	0,0000
AR(2)			
MA(1)			
MA(2)			
SAR(1)			
SAR(2)			
SMA(1)	-0,9361	0,1300	0,0000
SMA(2)	0,3536	0,1419	0,0127

FONTE: O autor (2018).

Observando o GRÁFICO A.5.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

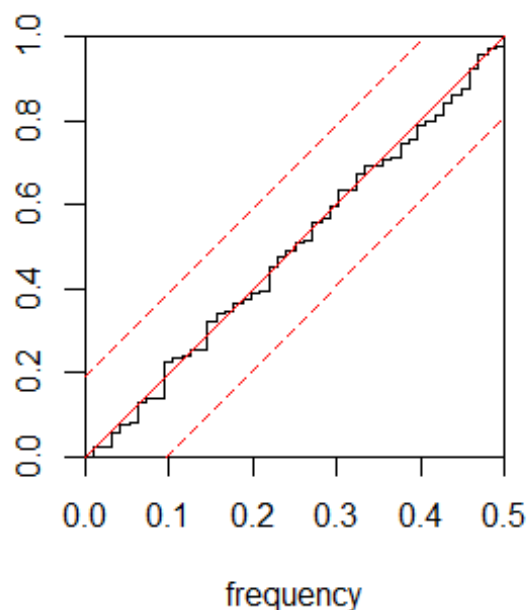
GRÁFICO A.5.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMBÉ



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.5.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

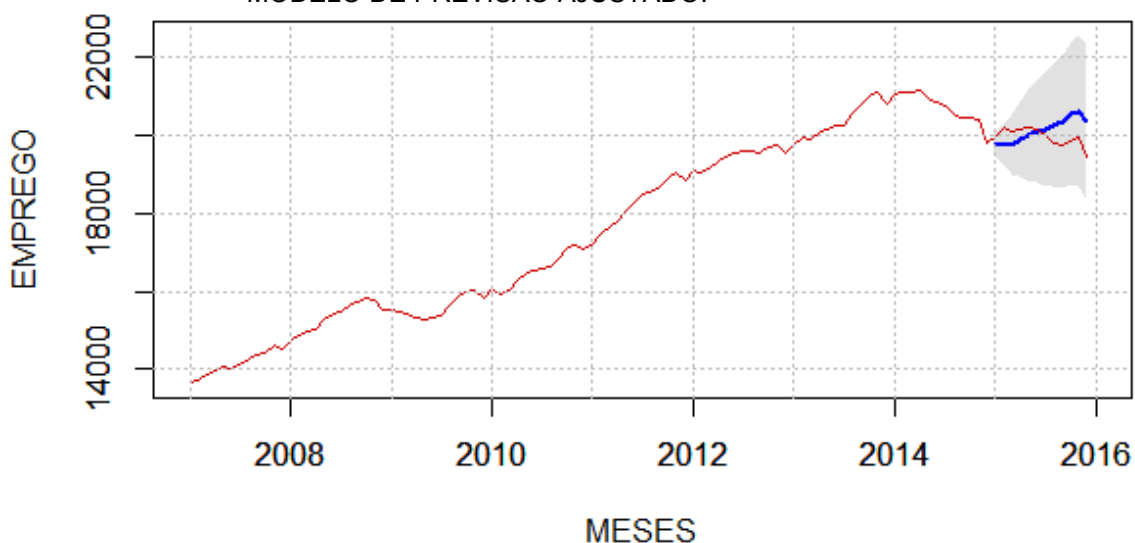
GRÁFICO A.5.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMBÉ



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.5.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.5.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CAMBÉ PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 6 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE CAMPO LARGO

Apresenta-se, abaixo, a TABELA A.6.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de Campo Largo. Observa-se que todos os coeficientes são significativos a um nível de significância de 0,5%.

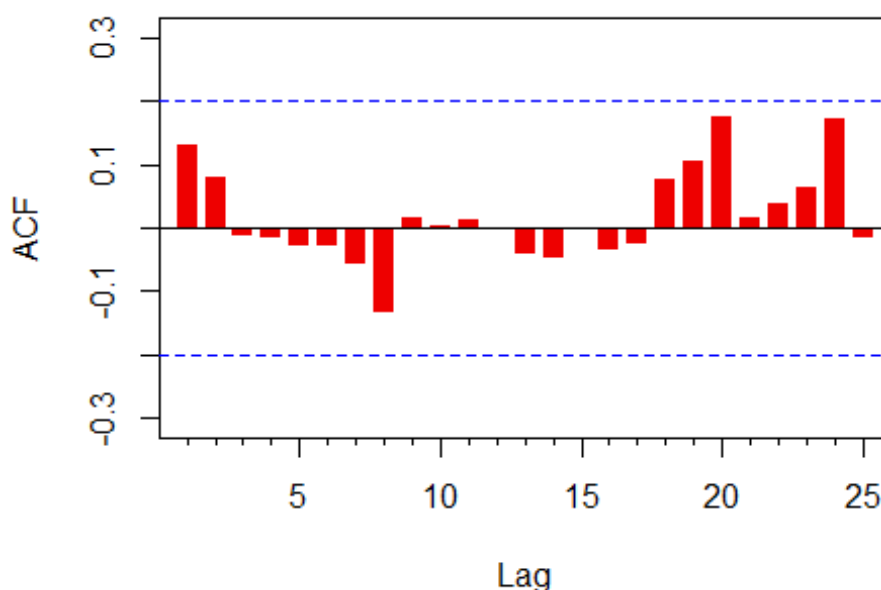
TABELA A.6.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMPO LARGO

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)			
AR(2)			
MA(1)	-1,0000	0,0166	0,0000
MA(2)			
SAR(1)	0,9818	0,0543	0,0000
SAR(2)			
SMA(1)	-0,8425	0,2380	0,0004
SMA(2)			

FONTE: O autor (2018).

Observando o GRÁFICO da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

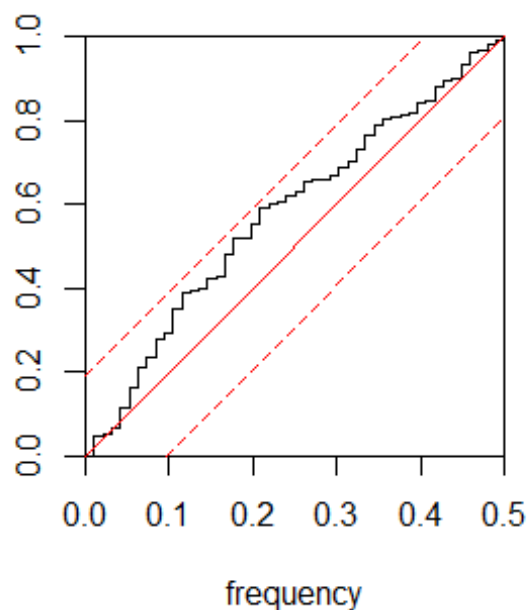
GRÁFICO A.6.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMPO LARGO



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.6.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

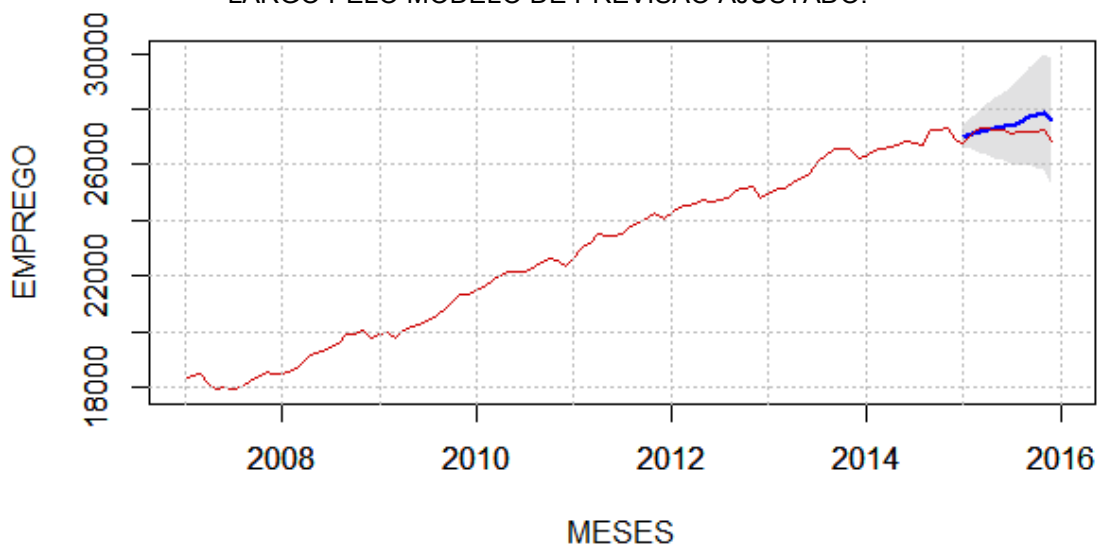
GRÁFICO A.6.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CAMPO LARGO



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.6.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.6.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CAMPO LARGO PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 7 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE CARAMBÉI

Apresenta-se, abaixo, a TABELA A.7.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de Carambéi. Observa-se que todos os coeficientes são significativos a um nível de significância de 0,5%.

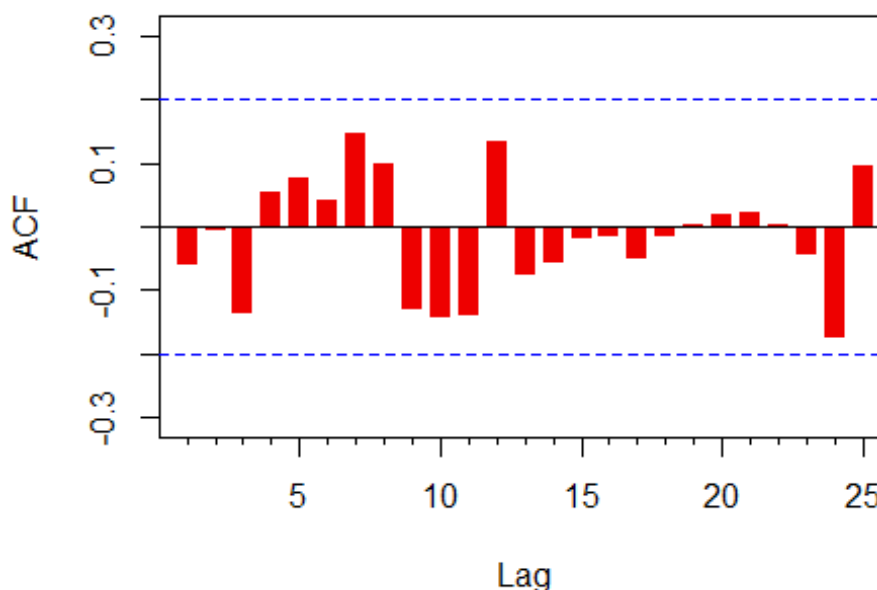
TABELA A.7.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CARAMBÉI

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	1,4658	0,1012	0,0000
AR(2)	-0,5121	0,1059	0,0000
MA(1)			
MA(2)			
SAR(1)	-0,616	0,101	0,0000
SAR(2)			
SMA(1)	-0,9994	0,2294	0,0000
SMA(2)			

FONTE: O autor (2018).

Observando o GRÁFICO A.7.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

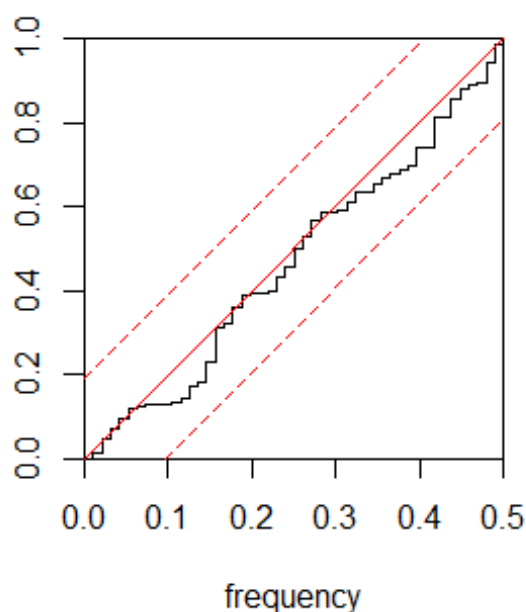
GRÁFICO A.7.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CARAMBÉI



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.7.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

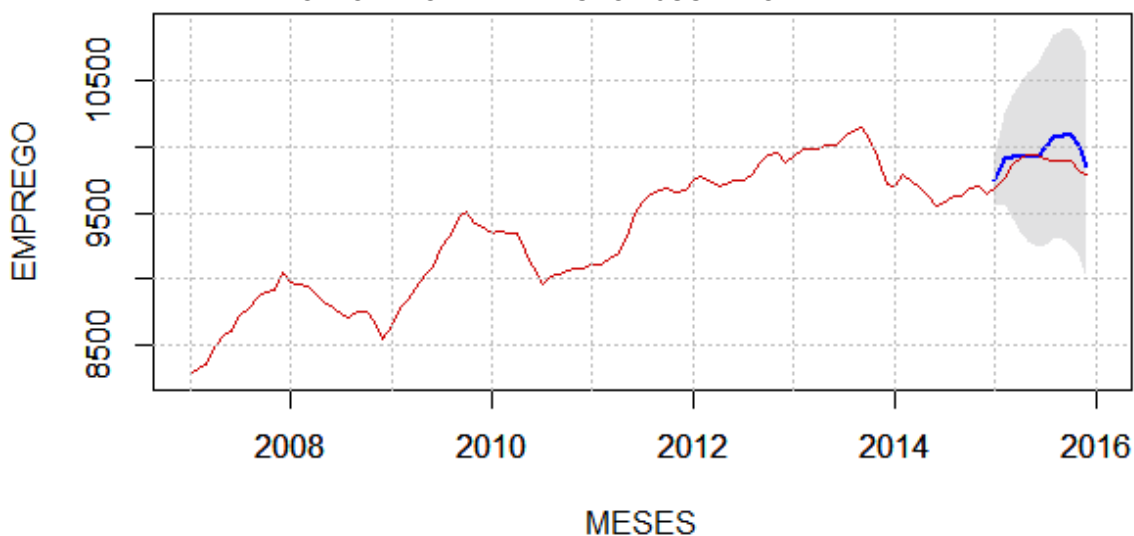
GRÁFICO A.7.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE CARAMBEÍ



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.7.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.7.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE CARAMBEÍ PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 8 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE COLOMBO

Apresenta-se, abaixo, a TABELA A.8.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de Colombo. Observa-se que todos os coeficientes são significativos a um nível de significância de 5%.

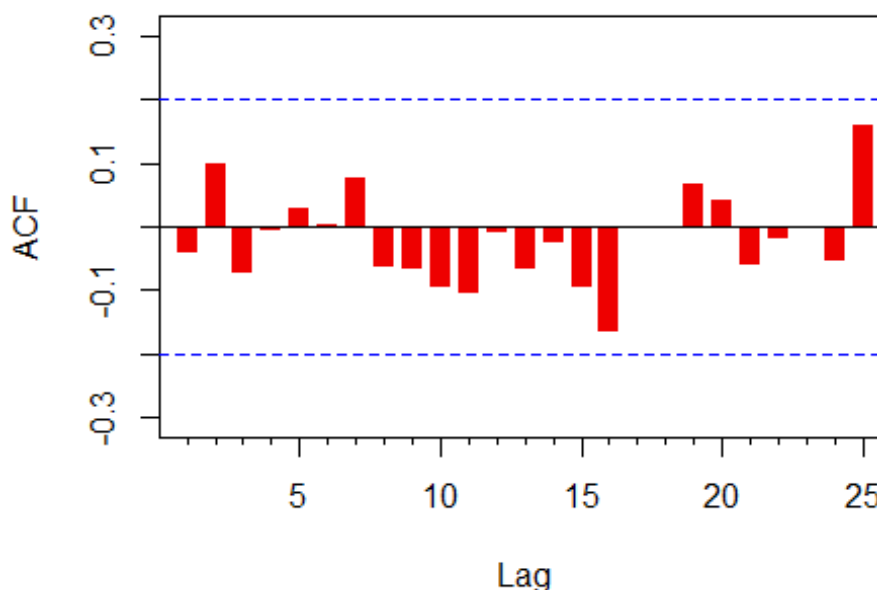
TABELA A.8.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE COLOMBO

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	0,2349	0,1042	0,0241
AR(2)			
MA(1)	-0,9730	0,0209	0,0000
MA(2)			
SAR(1)	0,9584	0,0532	0,0000
SAR(2)			
SMA(1)	-0,7367	0,1686	0,0000
SMA(2)			

FONTE: O autor (2018).

Observando o GRÁFICO A.8.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

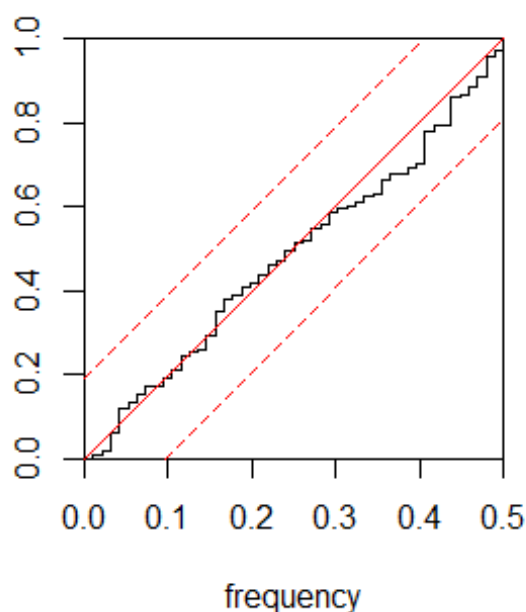
GRÁFICO A.8.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE COLOMBO



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.8.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

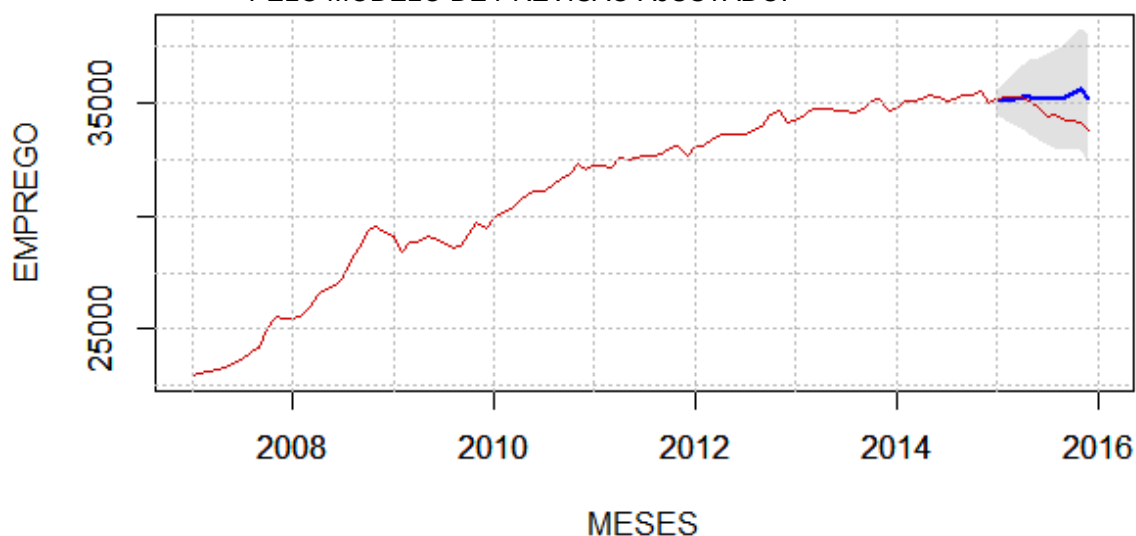
GRÁFICO A.8.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE COLOMBO



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.8.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.8.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE COLOMBO PELO MODELO DE PREVISÃO AJUSTADO.

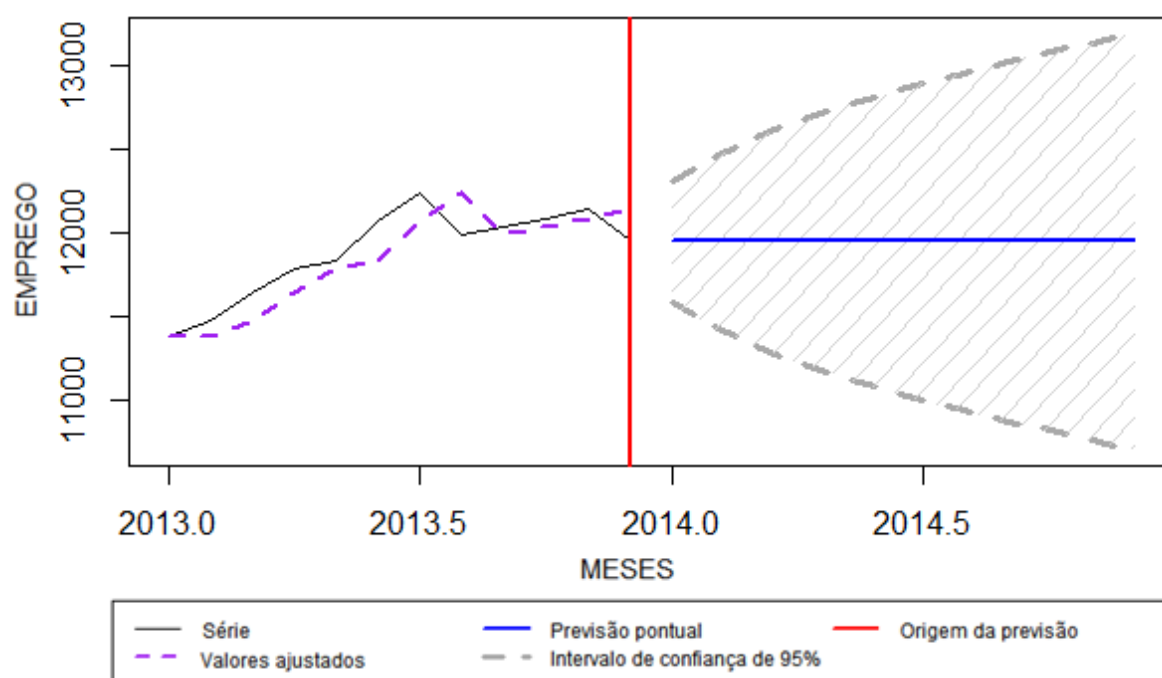


FONTE: O autor (2018).

APÊNDICE 9 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE FAZENDA RIO GRANDE

Apresenta-se, abaixo, o GRÁFICO A.9.1 do modelo de médias móveis simples estimado para o município de Fazenda Rio Grande. O gráfico contém os valores observados até o ano de 2014 e a projeção para o ano de 2015 a um nível de 95% de confiança.

GRÁFICO A.9.1 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE FAZENDA RIO GRANDE PELO MODELO MMS(1)



FONTE: O autor (2018).

APÊNDICE 10 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE IBIPORÃ

Apresenta-se, abaixo, a TABELA A.10.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arma()* para o município de Ibiporã. Observa-se que todos os coeficientes são significativos a um nível de significância de 0,5%.

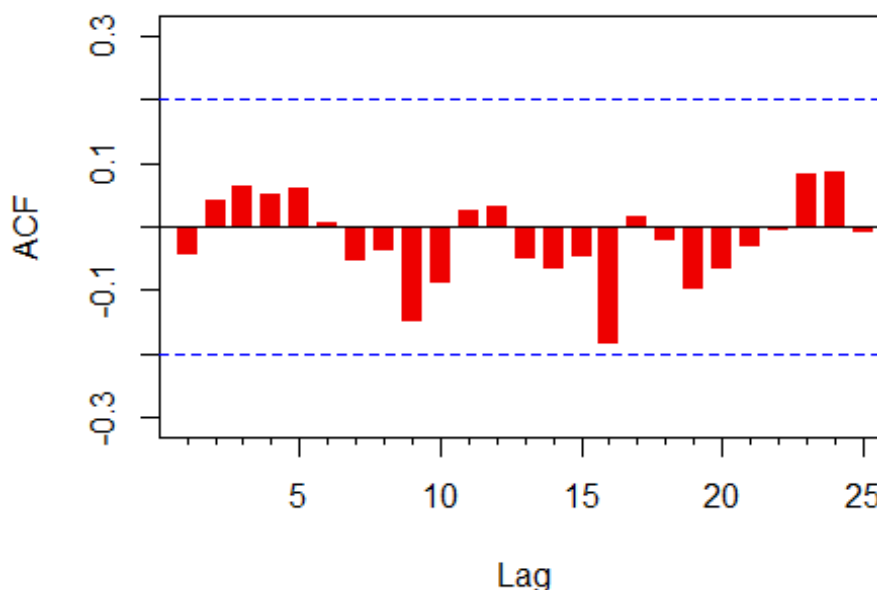
TABELA A.10.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE IBIPORÃ

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	-0,4388	0,1140	0,0001
AR(2)	0,6311	0,0693	0,0000
AR(3)	0,7995	0,1224	0,0000
MA(1)	1,9603	0,1289	0,0000
MA(2)	1,6932	0,2015	0,0000
MA(3)	0,6134	0,0935	0,0000
SAR(1)			
SAR(2)			

FONTE: O autor (2018).

Observando o GRÁFICO A.10.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

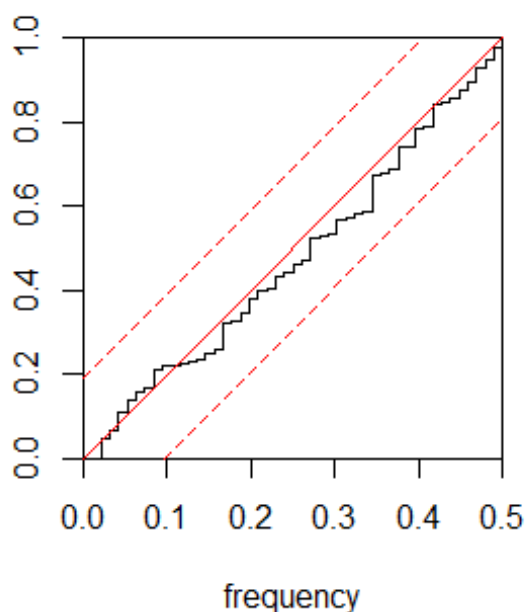
GRÁFICO A.10.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE IBIPORÃ



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.10.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

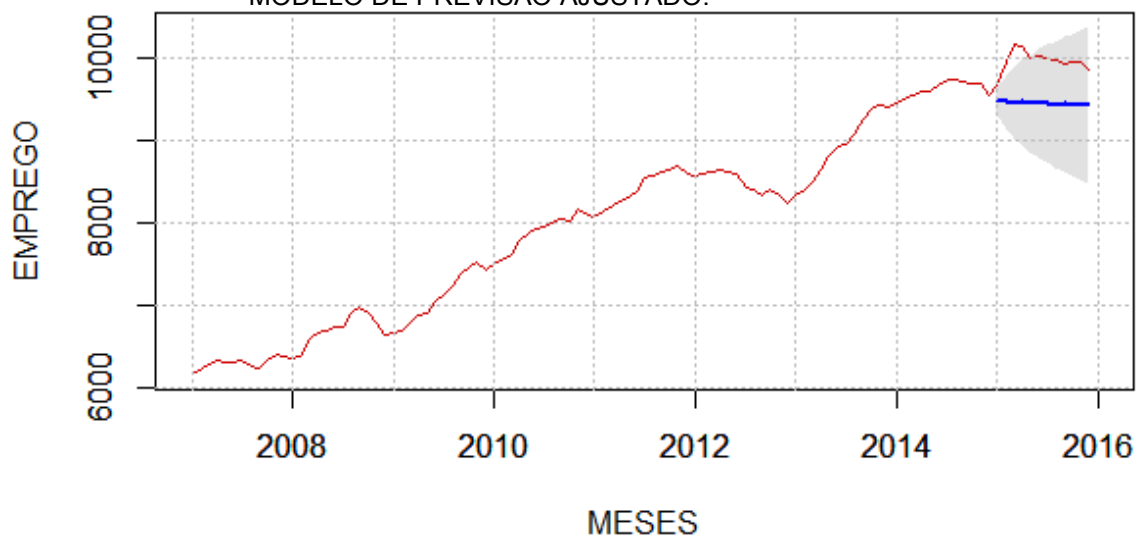
GRÁFICO A.10.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE IBIPORÃ



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.10.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.10.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE IBIPORÃ PELO MODELO DE PREVISÃO AJUSTADO.

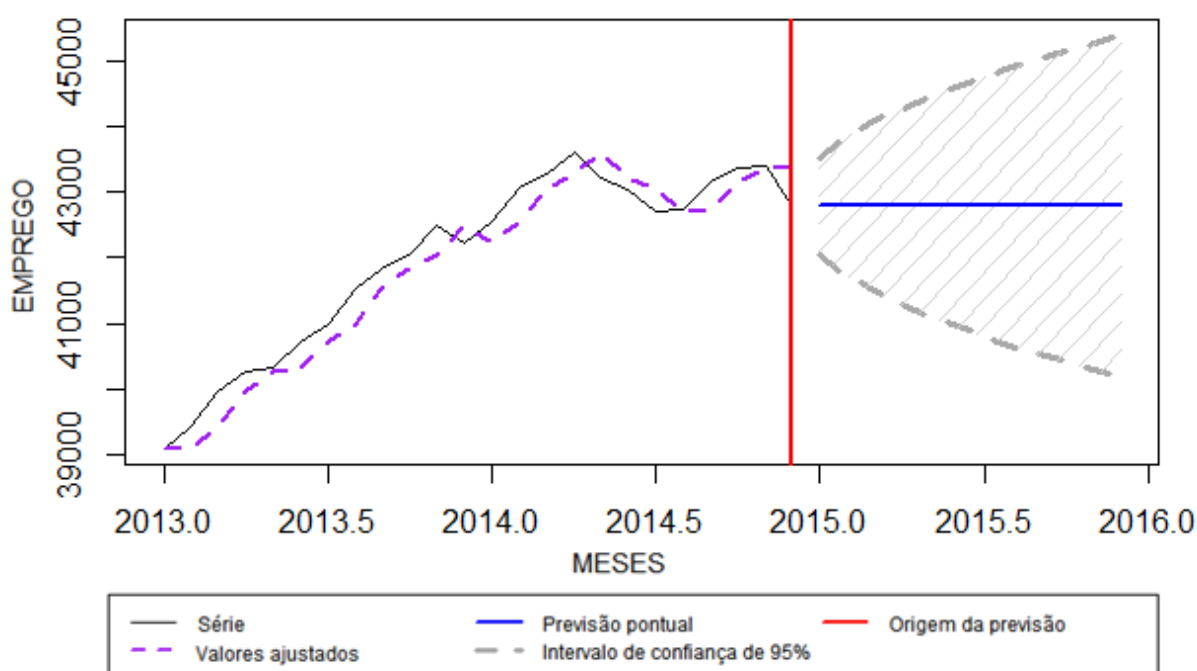


FONTE: O autor (2018).

APÊNDICE 11 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE PINHAIS

Apresenta-se, abaixo, o GRÁFICO A.11.1 do modelo de médias móveis simples estimado para o município de Pinhais. O gráfico contém os valores observados até o ano de 2014 e a projeção para o ano de 2015 a um nível de 95% de confiança.

GRÁFICO A.11.1 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE FAZENDA RIO GRANDE PELO MODELO MMS(1)



FONTE: O AUTOR (2018).

APÊNDICE 12 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE PIRAQUARA

Apresenta-se, abaixo, a TABELA A.12.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de Piraquara. Observa-se que todos os coeficientes são significativos a um nível de significância de 5%.

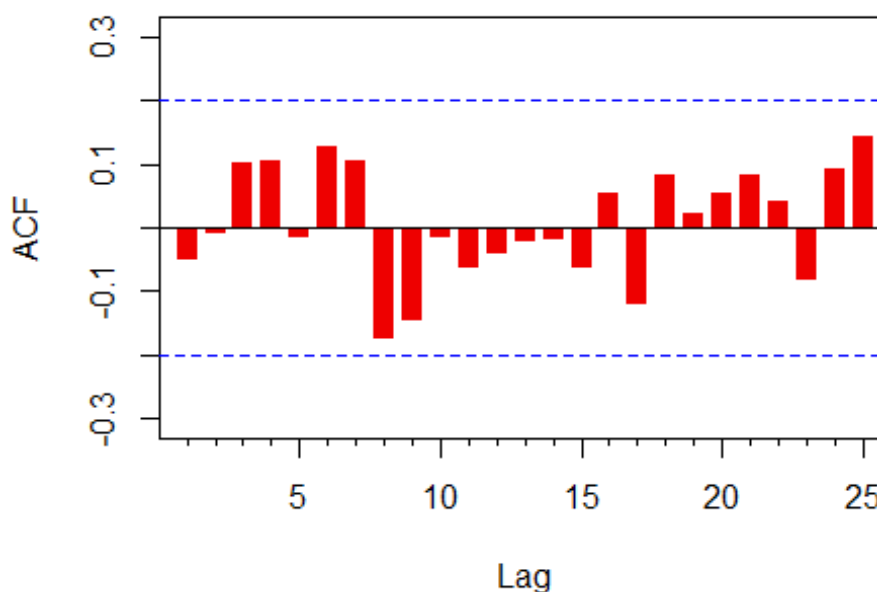
TABELA A.12.1 – COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE PIRAQUARA

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	0,4569	0,0846	0,0000
AR(2)	1,1127	0,0972	0,0000
AR(3)	0,2717	0,0819	0,0009
AR(4)	-0,8592	0,0685	0,0000
MA(1)	0,6752	0,1181	0,0000
MA(2)	-0,3226	0,1430	0,0241
MA(3)	-0,8610	0,1168	0,0000
SAR(1)	0,2641	0,1126	0,0190
INTERCEPTO	5656,0827	250,5244	0,0000

FONTE: O autor (2018).

Observando o GRÁFICO A.12.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

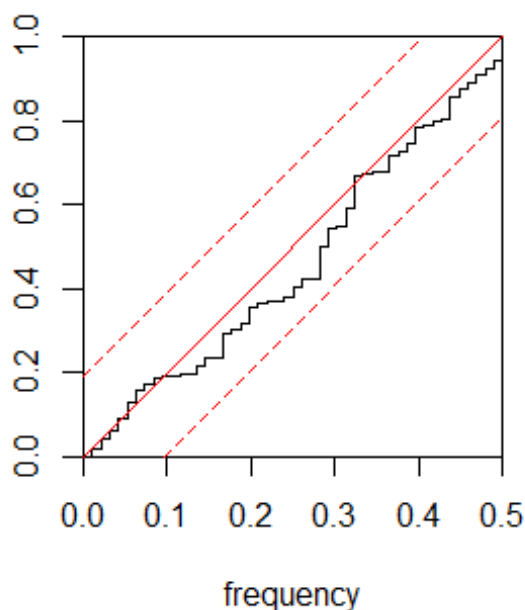
GRÁFICO A.12.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE PIRAQUARA



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.12.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

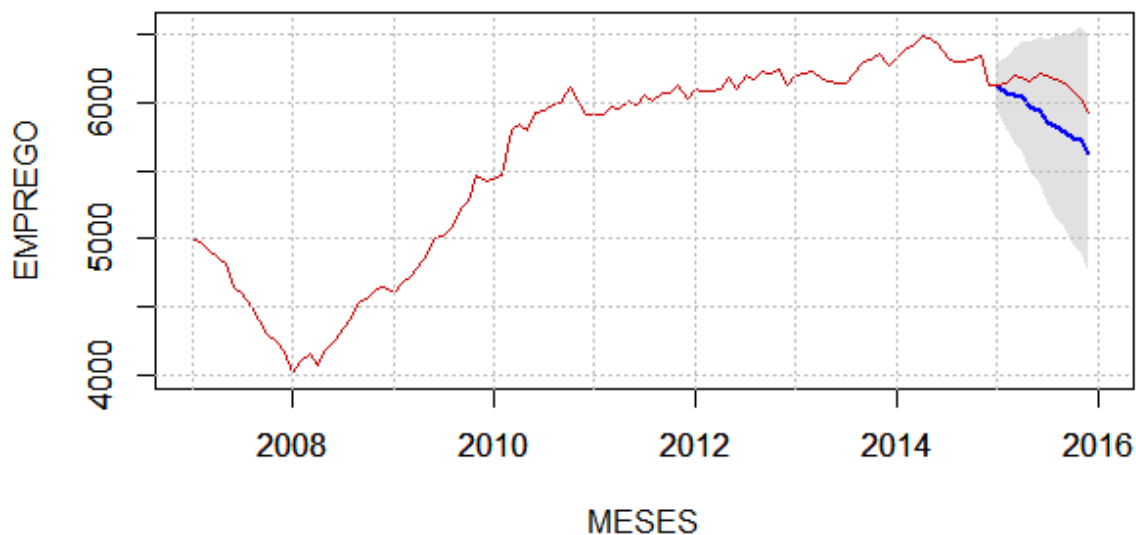
GRÁFICO A.12.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE PIRAQUARA



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.12.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.12.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE PIRAQUARA PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 13 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE ROLÂNDIA

Apresenta-se, abaixo, a TABELA A.13.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de Rolândia. Observa-se que todos os coeficientes são significativos a um nível de significância de 5%.

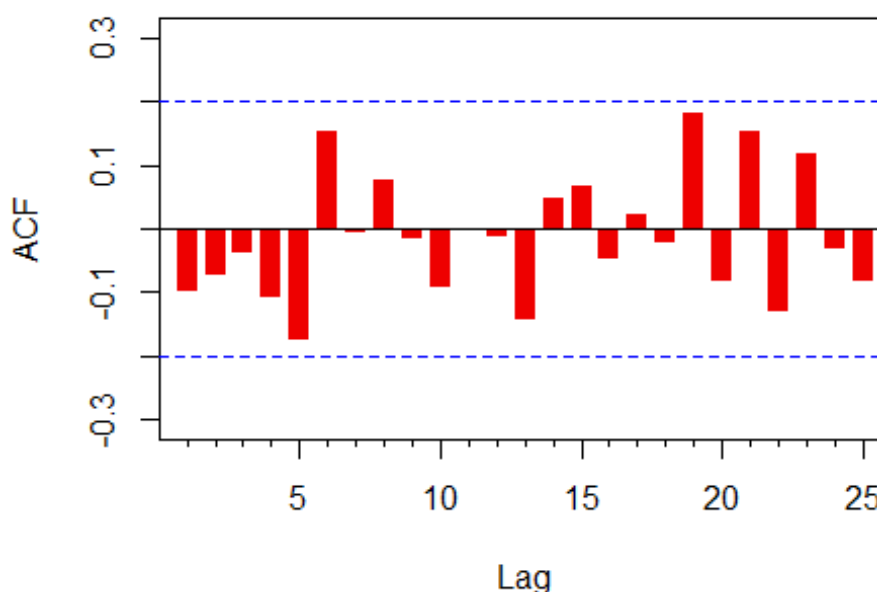
TABELA A.13.1– COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ROLÂNDIA

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	-0,8081	0,1190	0,0000
AR(2)	-0,5466	0,1399	0,0001
AR(3)	-0,4911	0,1303	0,0002
AR(4)	-0,2886	0,1203	0,0164
MA(1)	0,0000	0,0000	0,0000
MA(2)	0,0000	0,0000	0,0000
SAR(1)	-0,6609	0,1322	0,0000
SAR(2)	-0,3115	0,1583	0,0491
SMA(1)	-0,9467	0,3071	0,0021
SMA(2)	0,0000	0,0000	0,0000

FONTE: O autor (2018).

Observando o GRÁFICO A.13.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

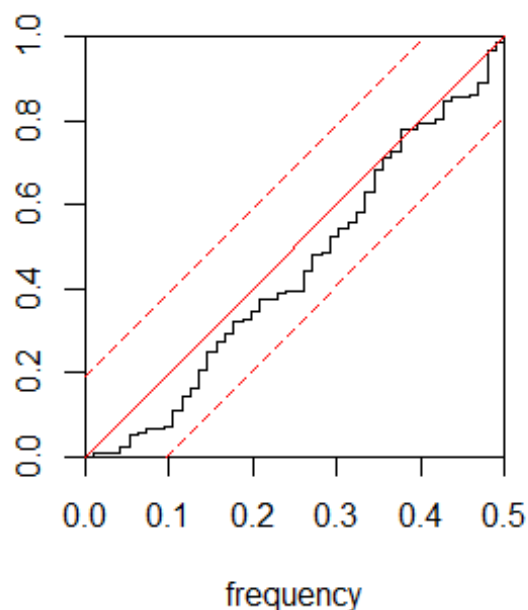
GRÁFICO A.13.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ROLÂNDIA



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.13.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

GRÁFICO A.13.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE ROLÂNDIA



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.13.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.13.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE ROLÂNDIA PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 14 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS

Apresenta-se, abaixo, a TABELA A.14.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arima()* para o município de São José dos Pinhais. Observa-se que todos os coeficientes são significativos a um nível de significância de 1%.

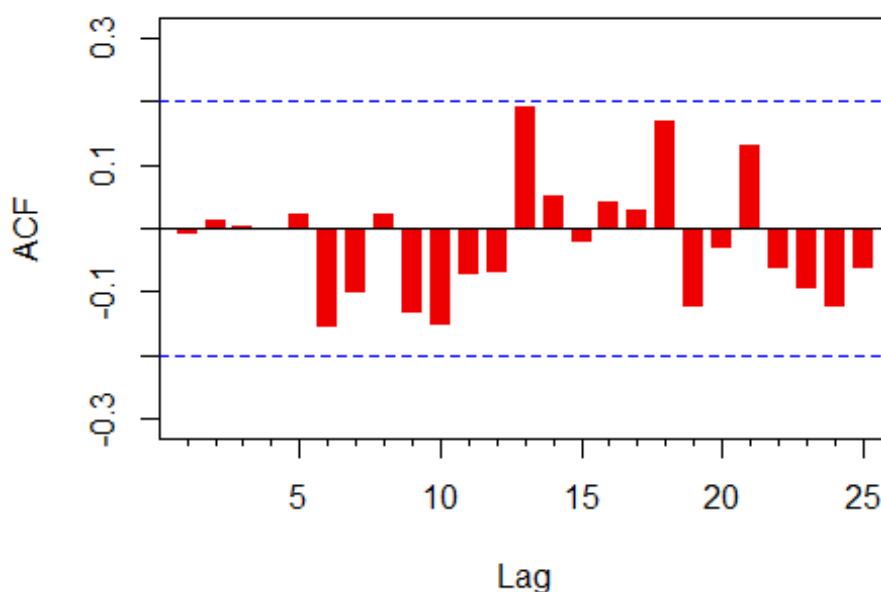
TABELA A.14.1– COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)	0,4015	0,1051	0,0001
AR(2)	0,2813	0,1029	0,0063
MA(1)	-0,9725	0,0252	0,0000
MA(2)			
SAR(1)	0,4084	0,0947	0,0000
SAR(2)	0,3634	0,1040	0,0000
SMA(1)			
SMA(2)			

FONTE: O autor (2018).

Observando o GRÁFICO A.14.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

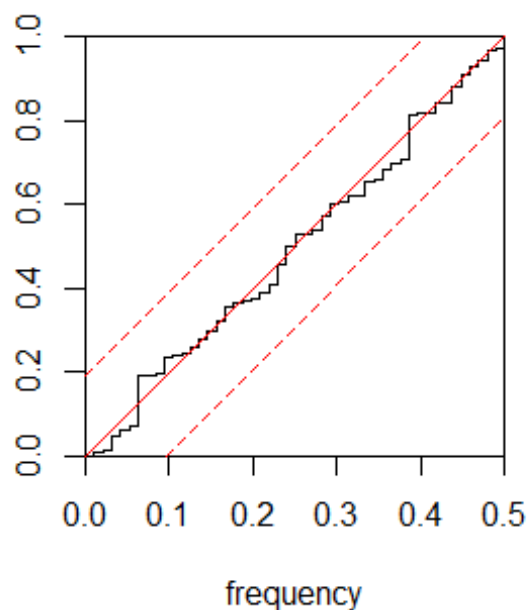
GRÁFICO A.14.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.14.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

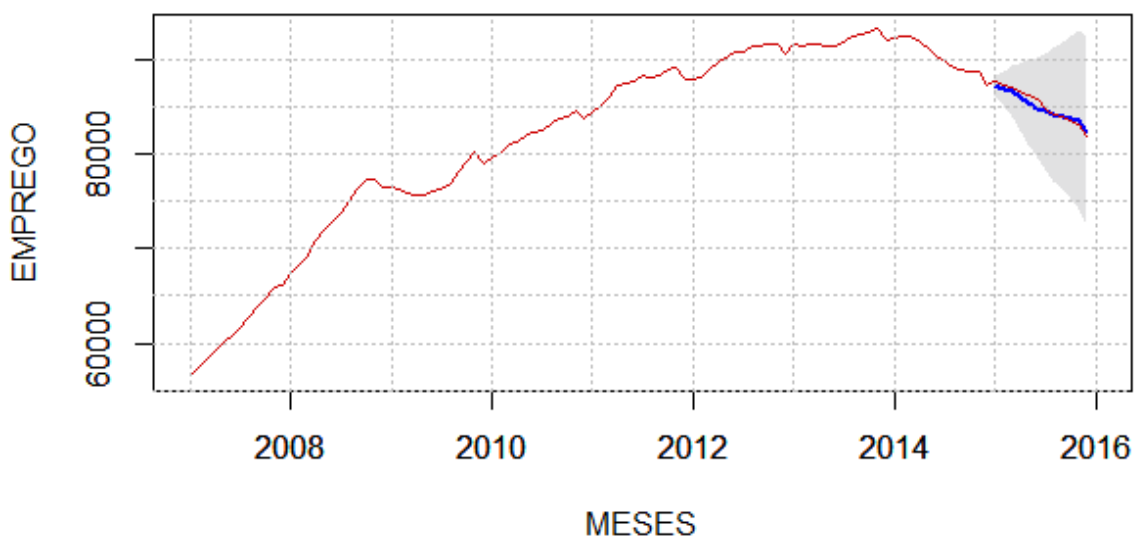
GRÁFICO A.14.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.14.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.14.3. – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE SÃO JOSÉ DOS PINHAIS PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 15 – PRINCIPAIS RESULTADOS DO MODELO DE PREVISÃO PARA O MUNICÍPIO DE SARANDI

Apresenta-se, abaixo, a TABELA A.15.1 dos coeficientes estimados para o modelo de previsão obtido pela função *best.arma()* para o município de Sarandi. Observa-se que todos os coeficientes são significativos a um nível de significância de 0,5%.

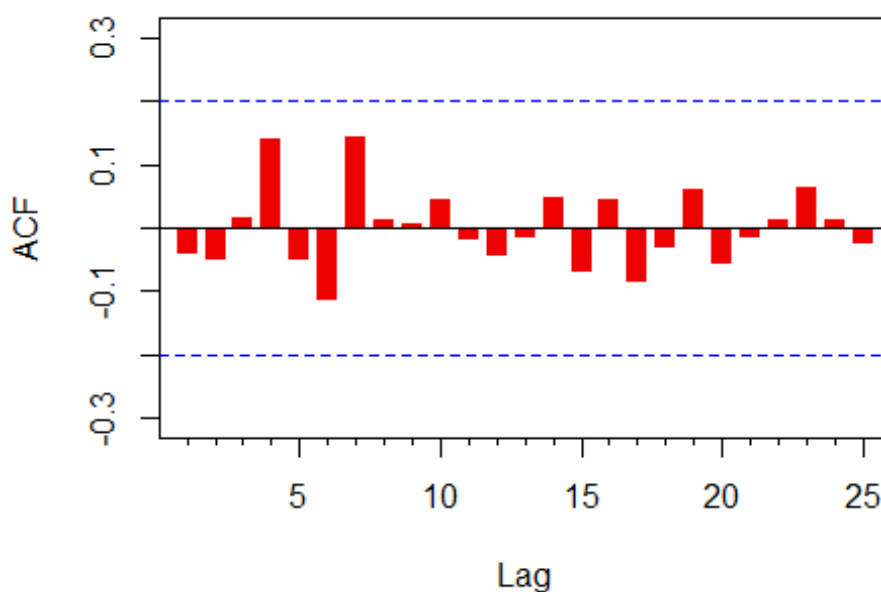
TABELA A.15.1– COEFICIENTES ESTIMADOS, ERRO PADRÃO E P-VALOR DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SARANDI

PARÂMETRO	COEFICIENTE ESTIMADO	ERRO PADRÃO	P-VALOR
AR(1)			
AR(2)			
MA(1)	-0,4097	0,1103	0,0002
MA(2)	-0,3660	0,1108	0,0009
SAR(1)			
SAR(2)			
SMA(1)	-0,7113	0,1320	0,0000
SMA(2)			

FONTE: O autor (2018).

Observando o GRÁFICO A.15.1 da ACF dos resíduos, abaixo, nota-se que todas as autocorrelações estão dentro do intervalo de confiança para a hipótese de nulidade do parâmetro $\hat{\rho}_k$.

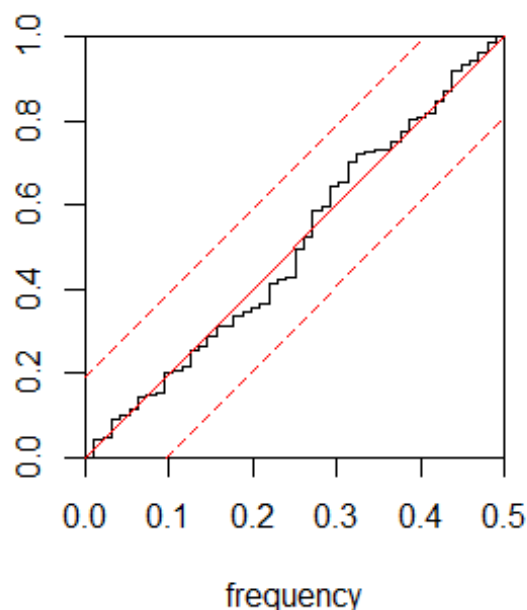
GRÁFICO A.15.1 – AUTOCORRELAÇÃO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SARANDI



FONTE: O autor (2018).

O periodograma acumulado dos resíduos é apresentado no GRÁFICO A.15.2. Devido ao fato do gráfico da série dos resíduos estimados (\hat{a}_t) se encontrar dentro das linhas limites, significa que \hat{a}_t é ruído branco.

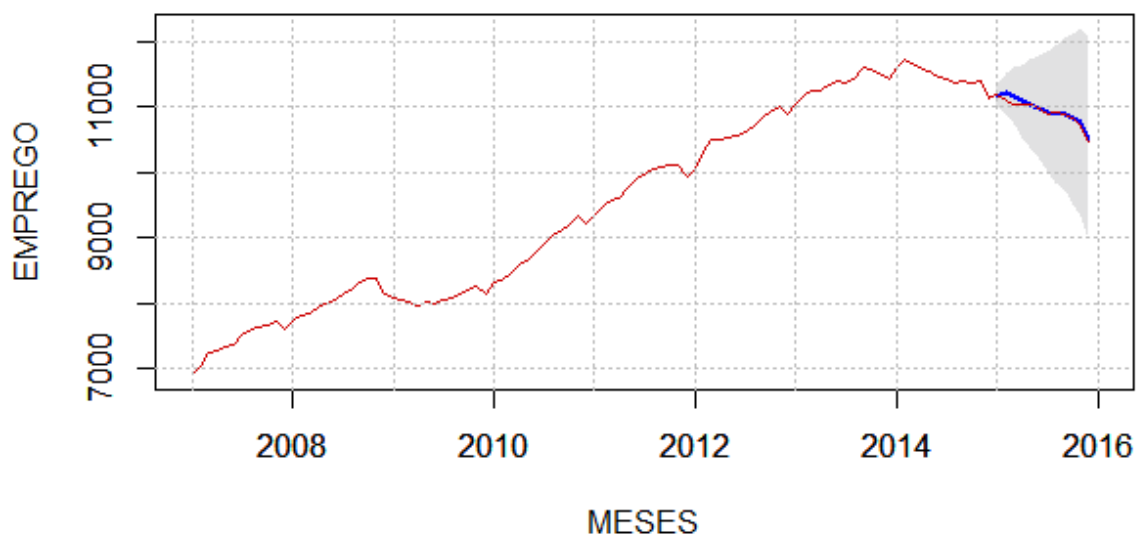
GRÁFICO A.15.2 – PERIODOGRAMA ACUMULADO DOS RESÍDUOS DO MODELO DE PREVISÃO AJUSTADO PARA O MUNICÍPIO DE SARANDI



FONTE: O autor (2018).

Por fim, apresenta-se o GRÁFICO A.15.3 de previsão do emprego celetista para o ano de 2015 e os dados obtidos pela RAIS para o município em questão. A previsão pontual é apresentada em azul e os valores observados em vermelho.

GRÁFICO A.15.3 – PREVISÃO DO EMPREGO CELETISTA PARA O MUNICÍPIO DE SARANDI PELO MODELO DE PREVISÃO AJUSTADO.



FONTE: O autor (2018).

APÊNDICE 16 – FUNÇÃO BEST.ARIMA() UTILIZANDO O SOFTWARE R CORE TEAM

```

#=====#
#===== FUNÇÃO BEST.ARIMA () =====#
#=====#

best.arima <- function(x, maxp=2, maxd=2, maxq=2, maxP=2, maxD=2, maxQ=2,
                      minp=0, mind=0, minq=0, minP=0, minD=0, minQ=0,
                      plot.modelos=FALSE) {
  library(tseries)
  library(forecast)

  ljung.n <- function(dados, lag.final, fitdf = 1) {
    pvalores <- matrix(NA, nrow = lag.final, ncol = 1)
    for (i in 1:lag.final) {
      pvalores[i] <- Box.test(dados,type = "Ljung-Box" , lag = i)$p.value
    }

    return(pvalores)
  }

  dados <- list(x[1:12],
               x[13:24],
               x[25:36],
               x[37:48],
               x[49:60],
               x[61:72],
               x[73:84],
               x[85:length(x)])

  KW = kruskal.test(dados) #testar sazonalidade
  if (KW$p.value > 0.05) {
    testes.TS=function(TS){
      phillips.perron = pp.test(TS)$p.value
      JB=jarque.bera.test(TS)$p.value

      dif=0
      while(JB<0.05 | phillips.perron>0.05){
        TS=diff(TS)
        phillips.perron = pp.test(TS)$p.value #Estacionariedade. H0:
        JB=jarque.bera.test(TS)$p.value

        dif=dif+1
      }

      cat("A SÉRIE FOI DIFERENCIADA",dif,"VEZES","\n")
      return(dif)
    }

    d=testes.TS(x)

    cont = 10000000000
    for (p in 0:maxp) {
      for (q in 0:maxq) {

```

```

try (a <- arima(x, order = c(p, d, q)))

if (plot.modelos==TRUE) {
  cat("ARIMA(", p, d, q, ")", "\n")
}

A = Acf(residuals(a), type = "correlation", lag.max = 25,
  plot = FALSE)
T = (A$acf < 2/sqrt(length(x)) & A$acf > -2/sqrt(length(x)))
T = table(T)
T = data.frame(T)

if (T[1, 2] == 1)
{
  p_valor_a = (1 - pnorm(abs(a$coef) / sqrt(diag(
    a$var.coef)))) * 2

  Pval = Pval = (p_valor_a > 0.05 | p_valor_a == "NaN")
  Pval = table(Pval)
  Pval = data.frame(Pval)
  p_valor_a = as.numeric(as.character(p_valor_a))

  #print(a)
  #print(p_valor_a)

  if (dim(Pval)[1] == 1) {
    if (p_valor_a[1] == "NaN"){p_valor_a[1] = 1.0}

    if (p_valor_a[1] < 0.05) {
      SW = shapiro.test(a$residuals)
      JB= jarque.bera.test(a$residuals)
      if (SW$p.value > 0.05 & JB$p.value > 0.05) {
        BPT = ljung.n(a$residuals, lag.final = 25)

        BP = (BPT > 0.05)
        BP = table(BP)
        BP = data.frame(BP)

        if (BP[1, 2] == 25) {
          if (a$aic < cont) {
            cont = a$aic
            b <- a
            pp <- p
            dd <- d
            qq <- q
            PP <- P
            DD <- D
            QQ <- Q

            arimaorder(b)
            cat("trocou para ARIMA(", p, d, q, ")", "\n")

            #cpgram(a$residuals, ci.col = "red");abline(0, 2, col = "red")
            if (p>2 | q>2 | P>2 | Q>2) {return(a)}
          }
        }
      }
    }
  }
}

```

```

    }
  }

  if(cont==10000000000) {print("NÃO FOI ENCONTRADO MODELO ALGUM")}
}

else{
  cont = 100000000000
  for (i in minQ:maxQ) {
    Q = i
    for (j in minD:maxD) {
      D = j
      for (k in minP:maxP) {
        P = k
        for (l in minq:maxq) {
          q = l
          for (m in mind:maxd) {
            d = m
            for (n in minp:maxp) {
              p = n

              try (a <- arima(x, order = c(p, d, q),
                            seasonal = list(order = c(P, D, Q), period = 12)
              ))

              if (plot.modelos==TRUE) {
                cat("ARIMA(", p, d, q, ")X(", P, D, Q, ")", "\n")
              }

              A = Acf(residuals(a), type = "correlation", lag.max = 25,
                      plot = FALSE)
              T = (A$acf < 2/sqrt(length(x)) & A$acf > -2/sqrt(length(x)))
              T = table(T)
              T = data.frame(T)

              if (T[1, 2] == 1)
              {
                p_valor_a = (1 - pnorm(abs(a$coef) / sqrt(diag(
                  a$var.coef
                )))) * 2

                Pval = Pval = (p_valor_a > 0.05 | p_valor_a == "NaN")
                Pval = table(Pval)
                Pval = data.frame(Pval)
                p_valor_a = as.numeric(as.character(p_valor_a))

                #print(a)
                #print(p_valor_a)

                if (dim(Pval)[1] == 1) {
                  if (p_valor_a[1] == "NaN"){p_valor_a[1] = 1.0}

                  if (p_valor_a[1] < 0.05) {
                    #SW = shapiro.test(a$residuals)
                    JB= jarque.bera.test(a$residuals)
                    if (#SW$p.value > 0.05 &
                        JB$p.value > 0.05) {
                      BPT = ljung.n(a$residuals, lag.final = 25)

```

